

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:51:25 ; Search time 40 Seconds  
(without alignments)  
1815.540 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKADPTNNDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

23: /SIDS2/gcgcdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2896	100.0	545	22 AAB49700	Small round struct
2	1979	68.3	544	22 AAB49703	Small round struct
3	1977.5	68.3	546	22 AAB49702	Small round struct
4	1950	67.3	530	15 AAR50972	Norwalk virus str
5	1938	66.9	530	22 AAB49701	Small round struct
6	1900	65.6	530	15 AAR57091	Small round virus
7	1225	42.3	540	22 AAB49706	Small round struct
8	1178.5	40.7	548	22 AAB49705	Small round struct
9	1174.5	40.6	535	22 AAB49707	Small round struct
10	1172.5	40.5	542	22 AAB49708	Small round struct

11	1157.5	40.0	539	22 AAB49704	Small round struct
12	1157	40.0	548	22 AAU91272	Norwalk virus asso
13	1154.5	39.9	550	22 AAB49709	Small round struct
14	1013	35.0	541	22 AAB49710	Small round struct
15	321	11.1	579	18 AAK08143	RUPV capsid protei
16	276	9.5	547	22 AAM50108	Feline calicivirus
17	276	9.5	671	22 AAM50107	Feline calicivirus
18	272.5	9.4	669	22 AAB67461	Amino acid sequenc
19	272	9.4	623	22 AAB47044	Feline Calicivirus
20	272	9.4	668	22 AAB67462	Amino acid sequenc
21	269	9.3	623	22 AAB47043	Feline Calicivirus
22	269	9.3	668	12 AAR10686	Feline calicivirus
23	269	9.3	668	22 AAE04304	Feline calicivirus
24	262.5	9.1	622	22 AAB47045	Feline Calicivirus
25	154	5.3	2206	13 AAR22210	True type 3 poliov
26	147	5.1	40	23 AAU91273	Norwalk virus asso
27	136	4.7	40	23 AAU91274	Norwalk virus asso
28	132.5	4.6	2209	3 AAP20037	Sequence encoded b
29	125.5	4.3	2179	7 AAP60243	Sequence encoding
30	116.5	4.0	1707	22 ABG22165	Novel human diagn
31	112.5	3.9	2164	9 AAP80131	Peptides translate
32	112.5	3.9	2164	9 AAP81045	Sequence of the vi
33	110	3.8	1037	22 ABB71260	Drosophila melanog
34	110	3.8	1234	22 ABB68510	Drosophila melanog
35	110	3.8	2016	22 ABB63911	Drosophila melanog
36	107.5	3.7	2599	21 AAY75098	Neisseria meningit
37	106	3.7	1302	20 AAY42104	Human MEK1 protei
38	106	3.7	1302	20 AAY26235	Human MEK1 protei
39	106	3.7	1493	20 AAY42107	Murine MEK1 prote
40	106	3.7	1493	20 AAY26234	Murine MEK1 prote
41	106	3.7	1593	23 AAM48935	Murine MEK1-2. M
42	105.5	3.6	239	9 AAP81097	Sequence of rhinov
43	105	3.6	1559	21 AAG50494	Arabisopsis thalia
44	105	3.6	1596	21 AAG50493	Arabisopsis thalia
45	105	3.6	1752	21 AAG50492	Arabisopsis thalia

ALIGNMENTS

RESULT 1

AAB49700

ID AAB49700 standard; protein; 545 AA.

XX

AC AAB49700;

XX

DT 04-APR-2001 (first entry)

XX

DE Small round structured virus protein SEQ ID 1.

XX

KW Small round structured virus; SRSV; food poisoning.

XX

OS Small round structured virus.

XX

PN WO200079280-A1.

XX

PD 28-DEC-2000.

XX

PF 22-JUN-2000; 2000WO-JP04095.

XX

PR 22-JUN-1999; 95Jp-0175928.

XX

PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX

DR WPI: 2001-080848/09.

XX

DR N-PSDB; AAF29141.

XX

PT Kit for the detection and typing of small round-structured virus (SRSV)

PT strains for investigation of food poisoning outbreaks, contains

PT antibodies

XX Claim 1; Page 40-42; 84pp; Japanese.

PS This invention relates to a kit for the detection and typing of small

XX round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the

CC spread of infection and to examine the epidemiology of outbreaks.

XX

XX Sequence 545 AA;

DR Query Match 100.0%; Score 2896; DB 22; Length 545;

XX Best Local Similarity 100.0%; Pred. No. 3.1e-250;

XX Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDAPTNDGTSAGOLVPEANTAEPTISMEPVAGATAAATAGOVNMIDPWIMNNY 60

Db 1 MMASKDAPTNDGTSAGOLVPEANTAEPTISMEPVAGATAAATAGOVNMIDPWIMNNY 60

QY 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQWNGVGNMKVKVLLAGNAFTA 120

Db 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQWNGVGNMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180

Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180

QY 181 LVCMLYTLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPLN 240

Db 181 LVCMLYTLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPLN 240

QY 241 TLNSRVPSLIKSMVSRDHQVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

Db 241 TLNSRVPSLIKSMVSRDHQVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

QY 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTTQNTGDIKQINVKQESAF 360

Db 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTTQNTGDIKQINVKQESAF 360

QY 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTAAQLAPP 420

Db 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTAAQLAPP 420

QY 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNECAPTRGEAALLHYLDPD 480

Db 421 IYPGFGGAIIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNECAPTRGEAALLHYLDPD 480

QY 481 THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540

Db 481 THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540

QY 541 GIRRS 545

Db 541 GIRRS 545

RESULT 2

AAAB49703

ID AAB49703 standard; protein: 544 AA.

XX

XX AAB49703;

AC

XX

XX 04-APR-2001 (first entry)

DT

DE Small round structured virus protein SEQ ID 4.

XX

XX Small round structured virus; SRSV; food poisoning.

XX

XX Small round structured virus.

OS

XX

PN WO200007940-A1.

XX 28-DEC-2000.

XX

XX 22-JUN-2000; 2000WO-JP04095.

PF

XX

XX 22-JUN-1999; 99JP-0175928.

PR

XX

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

PA

XX

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

PI

XX

XX WPI: 2001-080848/09.

DR N-PSDB; AAF29144.

DR

XX Kit for the detection and typing of small round-structured virus (SRSV)

XX strains for investigation of food poisoning outbreaks, contains

PT antibodies -

PT

XX

XX Claim 1; Page 47-49; 84pp; Japanese.

PS

XX This invention relates to a kit for the detection and typing of small

CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the

CC spread of infection and to examine the epidemiology of outbreaks.

CC

XX

XX Sequence 544 AA;

QY Query Match 68.3%; Score 1979; DB 22; Length 544;

Db Best Local Similarity 66.8%; Pred. No. 3.6e-168;

QY Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Db 1 MMASKDAPTNDGTSAGOLVPEANTAEPTISMEPVAGATAAATAGOVNMIDPWIMNNY 60

1 MMASKDAPTNDGTSAGOLVPEANTAEPTISMEPVAGATAAATAGOVNMIDPWIMNNY 60

QY 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQWNGVGNMKVKVLLAGNAFTA 120

Db 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQWNGVGNMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180

Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180

QY 181 LVCMLYTLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPLN 240

Db 181 LVCMLYTLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPLN 240

QY 241 TLNSRVPSLIKSMVSRDHQVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

Db 241 TLNSRVPSLIKSMVSRDHQVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300

QY 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTTQNTGDIKQINVKQES 358

Db 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTTQNTGDIKQINVKQES 358

QY 358

Db 358

QY 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTAAQLA 418

Db 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTAAQLA 418

QY 418

Db 418

QY 419 PPIYPGFGGAIIVFFMSDFPIAHGTNGLS- - -VPCITPQEFVTHFVNECAPTRGEAALL 474

Db 419 PPIYPGFGGAIIVFFMSDFPIAHGTNGLS- - -VPCITPQEFVTHFVNECAPTRGEAALL 474

QY 474

Db 474

QY 475 HYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVVFVSVSRFYQLKPVGTAG 534

Db 475 HYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVVFVSVSRFYQLKPVGTAG 534

QY 534

Db 534

QY 535 GPA-CRUGIRR 544

Db 535 GPA-CRUGIRR 544

Db 534 GPARGRLGVRR 544

RESULT 3  
AAB49702  
ID AAB49702 standard; protein; 546 AA.  
XX  
AC AAB49702;  
XX  
XX  
DT 04-APR-2001 (first entry)  
DE  
DE Small round structured virus protein SEQ ID 3.  
XX  
XX Small round structured virus; SRSV; food poisoning.  
KW  
XX Small round structured virus.  
OS  
XX WO200079280-A1.  
PN  
XX 28-DEC-2000.  
PD  
XX 22-JUN-2000; 2000WO-JP04095.  
XX  
PR 22-JUN-1999; 95JP-0175928.  
XX  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
PI  
XX WPI: 2001-080848/09.  
DR N-PSDB; AAF29143.  
XX  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies -  
XX  
PS Claim 1: Page 45-47; 84pp; Japanese.  
XX  
XX This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks.  
XX  
SQ Sequence 546 AA;

Query Match 68.3%; Score 1977.5; DB 22; Length 546;  
Best Local Similarity 66.6%; Pred. No. 4.9e-168;  
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;

Qy 1 MMASKDAPTNNDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVMNIDPWIMNY 60  
Db 1 MMASKDAPQSDAGSAGQLVPEYNTADPLPMEPVAGPTTAVATAGQVMNIDPWVNNF 60

Qy 61 VOAPQGETISPNPTPGDILFDLQGLPHLNPFLSHLAQYNGWGNMKVYKLLAGNAFTA 120  
Db 61 VQSPQGETISPNPTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRIILAGNAFSA 120

Qy 121 GKIIISCIPPGFAAQNISIAQTAMFPHVIADVRLIEVPLEDVRNVLFHNNNDAPTMR 180  
Db 121 GKIIIVCCVPPGPTSSSLTAQATLPHVIADVRLTEPIEMPLEDVRNVLHNTDNPQTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPPFIAGRVLTCPSPDFSFLFVPPNVEQTKPFSVNLPLN 240  
Db 181 LVCMLYTPLRGCGSNGSDFVAGRVLTAPSSDFSLFVPPTEQKTRAFVTPNPLQ 240

Qy 241 TLSNRSVPLSLIKMYSRHDGOMVQFNGRVTLDGOLGTTPTTSASOLCKIRGSVFHANG 300  
Db 241 TLSNRSFSLIOGMLSPDASQVQVQFNGRCLIDGLLGLTTPTATSGQLFRVRGKI--N98

Qy 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDMHMEASPT-TQFNTGDVVKOINVKQE-S 358  
Db 299 ARTLNLTEVDGKPFMAFSDPAPVGFDFGCKDWMHRISKTPTNTSSGDPMSRVSVOYTNVQ 358

Qy 359 AFAPHLGTIQADGLSDSVSYNTNNIAKLCHVSPVSDCHRGDVPWVLPYRGSTLTAEAAQLA 418  
Db 359 GFVPHLGSIQFDEVFNHPTG-DYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLA 417

Qy 419 PPIYPPGGEAIVFFMSDFPIAHGTNGLS----VPTIPQEFVTHFVNEQAPTREGAALL 474  
Db 418 PPVFPFGGEALVYFVSAFP---GPNRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALL 474

Qy 475 HYLDPDTHRLNGEFLKLYPEGFTVCVPNSSGTGPQTLPIGVFVSWSRFYQLKPVGTA 534  
Db 475 HYVDPDTHRLNGEFLKLYPGGYLTCPVNGVAGPQQLPLNGVLEFVSWSRFYQLKPVGTA 534

Qy 535 GPA-CRLGIRR 544  
Db 535 STARSRLGVRR 545

RESULT 4  
AAR50972  
ID AAR50972 standard; protein; 530 AA.  
XX  
AC AAR50972;  
XX  
DT 05-OCT-1994 (first entry)  
DE Norwalk virus strain 8FIia protein (encoded by ORF2).  
XX  
XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
KW seafood contamination; diagnostic assay; calicivirus;  
KW small round virus.  
XX  
OS Norwalk virus (strain 8FIia).  
XX  
PN WO9405700-A.  
XX  
PD 17-MAR-1994.  
XX  
PF 07-SEP-1993; 93WO-US08447.  
XX  
PR 07-SEP-1992; 92US-0941365.  
XX  
XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
PA  
XX Estes MK, Graham DY, Jiang X, Matson DO;  
PI  
XX WPI: 1994-101125/12.  
DR N-PSDB; AAQ56826.  
XX  
PT DNA from Norwalk and related viruses - used for preparing prods.  
PT for use in diagnostic assays, detection and vaccines for Norwalk  
PT and related viruses  
XX  
PS Claim 14; Page 68-70; 156pp; English.  
XX  
CC The Norwalk virus was isolated from stool samples from adult  
CC volunteers infected with safety tested Norwalk virus strain 8FIia.  
CC The coding sequence is useful for the design of probes for use in  
CC diagnostic assays for the Norwalk and related viruses.  
XX  
SQ Sequence 530 AA;

Query Match 67.3%; Score 1950; DB 15; Length 530;  
Best Local Similarity 67.0%; Pred. No. 1.3e-165;  
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

Qy 1 MMASKDAPTNNDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVMNIDPWIMNY 60  
Db 1 MMASKDATSSVDGAGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIMNF 60









```
PS Claim 1: Page 57-59; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
SQ Sequence 535 AA;

Query Match 40.6%; Score 1174.5; DB 22; Length 535;
Best Local Similarity 45.0%; Pred. No. 3.4e-96;
Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

Qy 1 MMASKDAPNTNDGTSAGOLVPEANTAEPTSMRPEVAGAAATAAGOVNMIDPWIMNNY 60
Db 1 MMASNDAAPSNDGAG---LVPEANN-ETWALEPVGASIAAPLTGQNNIIDPWIRLNF 56

Qy 61 VOAPOGFTISPNNTPGDILFDLQGLPHLPFLSLHAQMYNGWGNMKVKVLLAGNAFTA 120
Db 57 VOAPNGFTVSPRNSPGVLLNLELGPDLNPLVLAHLSRMVNGYAGGVEQVLLAGNAFTA 116

Qy 121 GKIIISCIPGFAAQNISIAQATWPHVIADVRLPEIEVPLEDVRNVLFH-NNDNAPTM 179
Db 117 GKLFAAAYPPHIFLEINSPGQITMFPHVIDVRLPEVLLPLPDVRNFFHYNQNEPRM 176

Qy 180 RLVCMLYPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFVSPNLP 239
Db 177 RLVMYLYPLRSNG--SGDDVFTVSCRVLTRPSDFDNLVLPPTLESKTPFTLPILTI 234

Qy 240 NTLNSRVPSLIKSMVSRDHQMVQFNGRVTLDGQLQGTTPTSASQLCKIRGSVFHA 299
Db 235 GELTNSRFPVIDELYTSPNESLVVQPGNGRCALDGLGQTTQLLPTAICSFGRINQV 294

Qy 300 GGNGY----NITELDGSYPH-AFESPAPIGFPD-----LGECDWHEASPTTOFWTG 346
Db 295 SGENHVNNQVNTINGTFDPTGDVPAPLPGTDFSGKLGFLVSQRD-HDNAC-----RSH 348

Qy 347 DVIKOINVKQESAFAPHLTGTQADGLSDVSVNTNNIAKLGWVSPVSDCHRGDVPWIPR 406
Db 349 DAVIATN--SAKFTPKGATQIGTWEEDDVHIHQPTKF---TPVGLFENEGFNQWTLPN 402

Qy 407 YGSLTEAAQLAPPITYPGFGEAIVFFMSDFPIAHGTNGLSVPCPTIPQEFVTHFVNEQAP 466
Db 403 YSGALTLNGLAPPVAPTPEGEQILFFRSHIPKGGVADPVIDCLLPQEWIQLHVESAP 462

Qy 467 TRGEAALLHYLDPDTHRNLEKFLYPEGFMTCVFNSSGTGPTPLPINGVVFVSVWSRFY 526
Db 463 SOSDVALIRFTNPDTGRVLFEAKLHRSGLITVA--NTGSRPIVVPANGYFRFDTWVNGFY 520

Qy 527 QLKPVGTAGPACRLGIRR 544
Db 521 SLAPMGTCN-----GRRR 533

RESULT 10
AAB49708
ID AAB49708 standard; protein: 542 AA.
XX
XX AAB49708;
AC AAB49708;
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 9.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX W0200079280-A1.
XX
```

```
PD 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP04095.
XX
PR 22-JUN-1999; 99JP-0175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29149.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies -
XX
XX Claim 1: Page 59-61; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks.
XX
XX Sequence 542 AA;

Query Match 40.5%; Score 1172.5; DB 22; Length 542;
Best Local Similarity 46.4%; Pred. No. 5.3e-96;
Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;

Qy 1 MMASKD-APTNDWDTSGAGOLVPEANTAEPTSMRPEVAGAAATAAGOVNMIDPWIMNN 59
Db 1 MMASNDAAPSN----DGAASLVPE-GINETMPLEPVAGASIAAPVAGOTNIDPWIRTN 55

Qy 60 YVQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSLHAQMYNGWGNMKVKVLLAGNAFT 119
Db 56 FVQAPNGEFTVSPRNSPGVLLNLELGPDLNPLVLAHLSRMVNGYAGGVEQVLLAGNAFT 115

Qy 120 AKTIISCIPGFAAQNISIAQATWPHVIADVRLPEIEVPLEDVRNVLFH-NNDNAPT 178
Db 116 AGKILFAAIPPNFLVDMISPAQITMLPLHIVDVRLPEITPLPDVRNVFYHFNQPOPR 175

Qy 179 MRLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFVSPNLP 238
Db 176 MRLVAMLYTPURSNG--SGDDVFTVSCRVLTRPTPDFEFYLVPPSVESKTKPFTLPILT 233

Qy 239 LNTLSNSRVPSLIKSMVSRDHQMVQFNGRVTLDGQLQGTTPTSASQLCKIRGSVFHA 298
Db 234 ISELNTRFPIEQLYTAPNETNVVQCGRCCTLDGELQGTQLLSAVCFLOQRTVAD 293

Qy 299 NGGN-GYNLTEL---DGSYPH-AFESPAPIGFPDLDGECDWHEASPTTOFWTGDV--IKQ 351
Db 294 NGDNWDONLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAK 353

Qy 352 INVKQESA-FAPHLGCTIQADGLSDVS--VNTNMIKLGWVSPVSDCHRGDVPWIPRYG 408
Db 354 IYISTSGKFTPKIGSI---GLHSITEHVHPNQQRSTPVGVAVD-ENTPFOQWVLPHYA 409

Qy 409 STLTEAAQLAPPITYPGFGEAIVFFMSDFPIAHGTNGLS--VPCTIPQEFVTHFVNEQAP 466
Db 410 GSLALNTNLAPAVAPTPEGEQILFFRSRVPVQVQLOGQDAFIDCLLPQEWNVHVFQEAAP 469

Qy 467 TRGEAALLHYLDPDTHRNLEKFLYPEGFMTCVFNSSGTGPTPLPINGVVFVSVWSRFY 526
Db 470 SOADVALLIRYVNPDTGRVLFEAKLHRSGLFITV--SHTGAYPLVVPNGHGRFDSWVNGFY 527

Qy 527 QLKPVGTAGPACRL 540
Db 528 SLAPMGTCNGRRRI 541
```





PR	22-JUN-1999;	99JP-0175928.	
XX	(NINA-)	JAPAN NAT INST INFECTIOUS DISEASES.	
PA	(DENK-)	DENKA SEIKEN KK.	
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
XX	WPI; 2001-080848/09.		
DR	N-PSDB; AAF29151.		
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies		
XX	Claim 1; Page 64-66; 84pp; Japanese.		
PS	This invention relates to a kit for the detection and typing of small		
XX	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRSV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks.		
XX	Sequence 541 AA;		
SQ	Query Match 35.0%; Score 1013; DB 22; Length 541;		
	Best Local Similarity 41.4%; Pred. No. 9.9e-82;		
	Matches 243; Conservative 81; Mismatches 155; Indels 108; Gaps 19;		
QY	1 MKWASKDAPNMDGTSGAGOLVPEANTAPISMEVAGATAAATAGQVNMIDPIMNMY 60		
Db	1 MKWASNDAPSSDGAAG---LPEINN-EVMPLEPVAGASLATPVVGQNIIDPWIRNF 56		
QY	61 VQAPQGEFTISPNTGDLDFDLQGLPHLNPFSLHIAQMYNGWGNKVKVLLAGNAFTA 120		
Db	57 VQAPAGEFTVSPNPGIILLDELGLDNLNPLYLAHARMYNGHAGMEVQIIVLAGNAFTA 116		
QY	121 GKIIISCIPGFAAQNISTAAQTMFHVITADVRLPVEVPLEDVRNLVLFH-NNDNAPTM 179		
Db	117 GKIIFAAIPGFPYENLSPQITMCPHVIIDVRLQLEPFLLPMPDIWNFFHYNGNDPKL 176		
QY	180 RLVCMLYPLRASGSSGTDFFVIAGRVLTCPSPDFSLFLVPNVVEQTKPFSVNLPL 239		
Db	177 RLVMALYPLRA--NNSGDDVFTVSCRVLTKPSPDFEFTFLVPPTVESKTKQFALPKI 234		
QY	240 NTLNSRVPSLIKSMYVRHGQWQFONGRVTLDGLOGTTTPTSASQLCKIRGVSFPHAN 299		
Db	235 SEMTNSRFPVPVDMYMTARNENQVQNGRVTLDGELGTTPLLAWNICKRGEVIAKN 294		
QY	300 GG-NGY---NLTELDGSPYHAFE-SPAPIGFPDLGECDDWHM-----EASPTQ--- 342		
Db	295 GDVRSYRMDMEITWTGTPIDPTEDTGPIGSPDFQGLFVGASQRKNQNPATRAHEA 354		
QY	343 -FNTG--DVIKQINVKQESAFAPHL-----GTQADGLSDVSVNTNMIK 384		
Db	355 IINTGGDHLCPQISSEIYLTSPNILRCTNPQLPQSLGRGTILI--RSDNGHCHDMVG- 411		
QY	385 LGWTSVPV-----SDGHRGDVDPVPIPRYGSGLTEAAQALPIYPGPF 426		
Db	412 ---TSPTPTTPWQWRKCRSGNCCSGHRYPV-PVVMNR----- 447		
QY	427 GEATVFFMSDFPIAHGTNGLSVPTCIPOEFVTFVNEQAPTRGEAALLHLYLDPDTHNLG 486		
Db	448 -----VTWIVLSHKSGFSSTRKLQ-----LNLRW-----LIREINPDTGRVLF 488		
QY	487 EFKLYPEGFMTCVNSSGTGQTLPIGCVFVFSWVSFRFLQKPVGT 533		
Db	489 EARLHKQGITVA--HTGDNPVIMPPNGYFRFEAWNQFYSLAPVGT 533		
RESULT 15			
AAW08143			

ID	AAW08143 standard; Protein; 579 AA.
XX	AAW08143;
AC	12-MAR-1997 (first entry)
XX	RHDV capsid protein.
DE	Canarypox virus; vaccinia virus; ALVAC; NVVAC; vector; attenuation;
XX	calicivirus; capsid; antigen; RHDV; vaccine; diagnosis; therapy.
XX	Rabbit haemorrhagic disease virus strain Saone et Loire.
OS	Key Location/Qualifiers
XX	Modified-site 45
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 281
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 308
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 369
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 393
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 430
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 474
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 481
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 502
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
XX	WO9639177-A1.
XX	12-DEC-1996.
XX	03-JUN-1996; 96WO-IB00721.
XX	06-JUN-1995; 95US-0471025.
XX	(VIRO-) VIROGENETICS CORP.
XX	Fischer L, Legros F, Paoletti E;
XX	WPI: 1997-042857/04.
DR	N-PSDB; AAT42749.
DR	Modified recombinant virus contg. exogenous calicivirus DNA - useful
PT	in vaccines and for in vitro proddn. of calicivirus antigens, for
PT	generation of therapeutic or diagnostic antibodies
XX	Example 12; Fig 12a; 150pp; English.
XX	The rabbit haemorrhagic disease virus (RHDV) capsid protein
CC	(AAW08143) amino acid sequence was deduced from the putative capsid
CC	gene (AAT42749) obtd. from RHDV strain Saone et Loire. The capsid
CC	protein can be expressed from novel recombinant, attenuated
CC	vaccinia virus and canarypox virus vectors, esp. based on NVVAC and
CC	ALVAC, useful as vaccines of improved safety. The protein may
CC	also be expressed in vitro and used to generate therapeutic or
CC	diagnostic antibodies.
XX	Sequence 579 AA;
SQ	

```
Query Match      11.1%; Score 321; DB 18; Length 579;
Best Local Similarity 28.3%; Pred. No. 9.2e-20;
Matches 132; Conservative 61; Mismatches 171; Indels 102; Gaps 21;

QY 12 MDGTSAGQLVPEANTAEPI$-----MEPVAGATAA-----ATAG----- 47
Db 1 MEGKARAAPQGEAAGTATTASVPGTTTGDGMDPGVATTSVITAENSASATATAGIGPPQ 60

QY 48 QVNMIDPWIMNNYVQAPQGEFTISPNNTPGDILFDLQGLNPLNPFSLHQAQMYNGWVGNM 107
Db 61 QVQOQETWRNFY-----NDVFTWSVADAGSILYTVQHSPPNNPFTAVLSQMYAGWAGM 117

QY 108 KVKVLLAGNAFTAGKIIISCIPGFAAQNISIA---QATMFPHVIADRVRLPIEVPLED 164
Db 118 QRFRIVAGSGVFGRLVAAVIPG-----IEIGPGLVRFQPHVVIDARSLPEVTTITMPD 172

QY 165 VRNVLFH-NNDNAPTWRLVCMLYTPL--RASGSSSGTDFFVIAGRVLTCPSPDFSLFLV 221
Db 173 LRPNMYHPTGDGLVPTLVLSVYNNLINPFGGSTS-----AIQVTVETTPSEDFEFVIR 227

QY 222 PPNVEQKTKPF$VPLNPLNLTLSNRVPSLIKSMVM$SRDHGMVQFO-----NGRVT 272
Db 228 APS--SKTVDSISFAGLLT-----PVLTCVGNDRN$NGQIVGLQPVPGG$TCNRHWN 279

QY 273 LDGLOQCTTPTSASOLCKIRGSVFHANGNGYNLT$EL---DGSPYHAFESP---APIG 324
Db 280 L$G$TYGWS$PRFADIDHRRGSASY-SGN$ATNVLQFWANAGS---AIDNPISQVAPDG 335

QY 325 FPDLGEC$DHMEASPTQFNTG$VIQ$INVKQESAFPHLGTIQADGLSDVSVNTN$IAK 384
Db 336 FPDMSFVPFNGPGIPAAGWVG$AIWNSN-----SGAPNVTTVQA-----YE 377

QY 385 LGWVSPVSDGHRGDVDPWIPRYG$TLTEAAQ-LAPPIYPPG$GEA 429
Db 378 LGPAT-----GAPCNLQP-----TTTSCAQTVAK$SIYAVVTGTA 412
```

Search completed: January 16, 2003, 09:54:50  
Job time : 43 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:54:10 ; Search time 16 Seconds  
(without alignments)  
1002.218 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep.\*  
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	285	9.8	626	4	US-09-590-020-7
2	272	9.4	623	4	US-09-590-020-4
3	269	9.3	623	4	US-09-590-020-2
4	262.5	9.1	622	4	US-09-590-020-6
5	149	5.1	2206	1	US-07-852-260-2
6	149	5.1	2206	2	US-08-461-503-2
7	149	5.1	2206	4	US-08-465-250-2
8	106	3.7	1302	4	US-09-423-890-2
9	106	3.7	1493	4	US-09-423-890-8
10	106	3.7	1593	4	US-08-628-829-4
11	106	3.7	2318	4	US-09-091-219-24
12	103.5	3.6	1170	4	US-09-749-588-2
13	102.5	3.5	588	1	US-08-460-860-4
14	102	3.5	1048	4	US-09-171-699-10
15	100	3.5	544	4	US-09-615-192A-349
16	98	3.4	1209	4	US-09-749-588-4
17	97	3.3	956	4	US-09-134-078-63
18	97	3.3	3443	2	US-08-416-603-2
19	95.5	3.3	202	2	US-08-416-603-10
20	95	3.3	2232	4	US-09-091-219-25
21	95	3.3	2247	4	US-09-091-219-2
22	94.5	3.3	2004	1	US-08-375-709-15
23	94.5	3.3	2004	1	US-08-752-929-15
24	94.5	3.3	2004	1	US-09-090-793-9
25	94	3.2	1036	4	US-09-206-942-73
26	94	3.2	1477	1	US-08-038-682-4
27	94	3.2	1477	1	US-08-302-832-4

28	94	3.2	1477	2	US-08-530-198-4	Sequence 4, Appl
29	94	3.2	1477	2	US-08-469-880-4	Sequence 4, Appl
30	94	3.2	1477	2	US-08-728-470-4	Sequence 4, Appl
31	94	3.2	1477	2	US-08-617-697-4	Sequence 4, Appl
32	94	3.2	1477	4	US-08-719-641-4	Sequence 4, Appl
33	94	3.2	1477	4	US-09-206-942-71	Sequence 71, Appl
34	93.5	3.2	907	3	US-08-783-774-2	Sequence 2, Appl
35	93.5	3.2	907	4	US-09-328-599A-1	Sequence 1, Appl
36	93.5	3.2	907	5	PCT-US95-04611A-19	Sequence 19, Appl
37	93	3.2	20	4	US-08-973-961-5	Sequence 5, Appl
38	93	3.2	1381	4	US-09-540-245A-16	Sequence 16, Appl
39	92.5	3.2	669	4	US-09-071-035-264	Sequence 264, App
40	92.5	3.2	1638	4	US-09-071-035-258	Sequence 258, App
41	92.5	3.2	1638	4	US-09-071-035-262	Sequence 262, App
42	92.5	3.2	1638	4	US-09-071-035-266	Sequence 266, App
43	92.5	3.2	2308	1	US-08-015-973-1	Sequence 1, Appl
44	92.5	3.2	2308	2	US-08-448-164-1	Sequence 1, Appl
45	92.5	3.2	2308	4	US-08-081-929-2	Sequence 2, Appl

## ALIGNMENTS

### RESULT 1

US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnius, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7

Query Match 9.8%; Score 285; DB 4; Length 626;

Best Local Similarity 23.8%; Pred. No. 4.7e-20;

Matches 144; Conservative 87; Mismatches 194; Indels 180; Gaps 32;

QY	13	DGTSGAGQLVPEANT-AEPISEMPVAGAAATAAGOVNMDPWNNVYQAQGEFTI- 70	
DB	82	DGSIITA-----PEOGTAVGGVIAEPAQMSAADMATCKSVDSW-----EAFESFH 128	
QY	71	-----SPNNTPGDILDLQGLPHLPFLSLHAQMYGNWGNMKVKVLLAGNAFTAGKIII 125	
DB	129	TSVMNSTSETQGIKFKQSLGSLPLNPLYSLHAKLYVAMWSGSEVRETSISGSGVGKLA 188	
QY	126	SCIPPGF-RAQNTISIAQAQMFPHVIADRVVLEPIEVPLEDVRNVLFHNNNDNATMRLVCM 184	
DB	189	IYVPPGVDVPQVSTSMIQ---YPHVLEFDAQREQVEPFIPLDRLSTLYHLMSDOTTSLVIM 245	
QY	185	LY-----TPLRASGSSSGTDPFVIAGRVLTCPSPDFELFLVPPNVQKTKPFVSNPLPLN 240	
DB	246	VYNDLNPVANDTNSGCG---IVT--VETKPGDPDFEHLKPPG-----S 285	
QY	241	TLSNRVPS-LI---KSMVSRDH-QGMVQFQNGRVTLQGLQGTTPTSASQLCKIRGSV 295	
DB	286	MLTHGSVPSPDLIPKSSSLWGNRHSWDITDF-----IIRPFV 322	
QY	296	PHANGNGYN-----LTELDCSPHYA-----FESPA-PIGPPDL---G 329	
DB	323	FOANRHFEDFNQETAGWSTFRFRPISVTITEQNGAKLIGVATDYIVPGIPDGWPDTPIG 382	

QY 306 LTELDGSYHAFESPA-PIGFPDLGEC-----WHMEASPTQF-----NT 345  
 Db 400 ITTAAG--YDAAEITVNTNFKSMYICGSLQRAWGDKKISNTAFIITAVRKNSIEPSNT 457  
 QY 346 GDVTKQINVKQESAFAPHLGTIQADGLSDYSVNTNITAKLGWSPVSDGHRGDVDP---- 401  
 Db 458 IDMTKLV-VYQDA---HVG-----EEVQTSIDITLALLGYTGIGEAIGSDRDKVRI 505  
 QY 402 WIPRYGSTLTEAAQLAPPYPPGFGGAIVFFMSDPFIAHGTNGLSVPCPTIPOEFVTHFV 461  
 Db 506 SVLPETGAR-----GGNHPIFYKNSIKLGVIRSIDV-----FN 539  
 QY 462 NEQAPTRGEAALLHYLDP-----DTHRNLEGEKLYPEGFMTCPVNSSSGTGPOTLP 511  
 Db 540 QILHTSRQLSNLYLLPPDSFAVYRIIDNSGWFIDIGSDGF-----SFVGSNLP 592

RESULT 3  
 US-09-590-020-2  
 ; Sequence 2, Application US/09590020  
 ; Patent No. 6355246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kruger, John M  
 ; APPLICANT: Maes, Roger K  
 ; APPLICANT: Vilnis, Alvars  
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
 ; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
 ; FILE REFERENCE: MSU 4.1-446  
 ; CURRENT APPLICATION NUMBER: US/09/590,020  
 ; CURRENT FILING DATE: 2000-06-08  
 ; PRIOR APPLICATION NUMBER: 60/138,484  
 ; PRIOR FILING DATE: 1999-06-10  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 623  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 US-09-590-020-2

Query Match 9.3%; Score 269; DB 4; Length 623;  
 Best Local Similarity 29.8%; Pred. No. 2e-18;  
 Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;  
 QY 7 DAPT-NMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWINNYYQAPQ 65  
 Db 72 DULPRLGDDGSIITPEQGTWVGVIAPSAQMSAAADMATGKSVSEW-----E 122  
 QY 66 GEFTI-----SPNTPGDIILDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAFT 119  
 Db 123 AFFSHTSVNMWSTSETQKILFKQSLGPNLNPYLSHLAKLYVANGSGVEVRESISGSGVF 182  
 QY 120 AGKIISCIPPGF-AAQNISIAQATMFPHVIADRVRLPIEVPLEDVNRVLFHNDNAPT 178  
 Db 183 GGLAAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEPVFAIPDLRSNLYHLMSDIDT 239  
 QY 179 MRLVCMY-----TPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPSV 234  
 Db 240 TSLVMVYNDLINPYANDTNSGCG---IVT--VETKPGDFKFKLLKPPG----- 284  
 QY 235 PNLPLNTLSNRVPS 249  
 Db 285 -----SMLTHGCVPS 294

RESULT 4  
 US-09-590-020-6  
 ; Sequence 6, Application US/09590020  
 ; Patent No. 6355246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kruger, John M  
 ; APPLICANT: Maes, Roger K  
 ; APPLICANT: Vilnis, Alvars

QY 330 EC-----DWHM-----EASPTTQFNTGDVIK-QINVK-----QESAPA 361  
 Db 383 ELIPAGDYAITNGTNDIITATCYDTADIIKNTNFRGMWICGSLQRAWGDKKISNTAF- 441  
 QY 362 PHLGTIOADGLSDSVN-TNMI-----AKLGWSPVSD-----GHRGDVDPWVI 404  
 Db 442 --ITTATLGDGNNKINPCNTIDQSKIVVFODAHVGKKAQTSDDTLALLGYTGIGEQAI- 498  
 QY 405 PRYGSTLTEAAQLA--PPIYPPGFGGAIVFFMSDRPFIAGHTNGLSVPCPTIPOEFVTHFVN 462  
 Db 499 ---GSDRDRVVRISTLPETGARG-GNHPIFYKNSIKLGVIRSIDV-----FNS 543  
 QY 463 EQAPTRGEAALLHYLDP-----DTHRNLEGEKLYPEGFMTCPVNSSSGTGPOTLP 512  
 Db 544 QILHTSRQLSNLYLLPPDSFAVYRIIDNSGWFIDIGSDGF--SFVGSVGFGLKLEFPL 601

QY 513 NGVEFV 517  
 Db 602 SASYM 606  
 RESULT 2  
 US-09-590-020-4  
 ; Sequence 4, Application US/09590020  
 ; Patent No. 6355246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kruger, John M  
 ; APPLICANT: Maes, Roger K  
 ; APPLICANT: Vilnis, Alvars  
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
 ; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
 ; FILE REFERENCE: MSU 4.1-446  
 ; CURRENT APPLICATION NUMBER: US/09/590,020  
 ; CURRENT FILING DATE: 2000-06-08  
 ; PRIOR APPLICATION NUMBER: 60/138,484  
 ; PRIOR FILING DATE: 1999-06-10  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 623  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 US-09-590-020-4

Query Match 9.4%; Score 272; DB 4; Length 623;  
 Best Local Similarity 22.2%; Pred. No. 9.7e-19;  
 Matches 133; Conservative 84; Mismatches 209; Indels 174; Gaps 27;  
 QY 7 DAPT-NMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWINNYYQAPQ 65  
 Db 72 DULPRLGDDGSIITPEQGTWVGVIAPSAQMSAAADMATGKSVSEW-----E 122  
 QY 66 GEFTI-----SPNTPGDIILDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAFT 119  
 Db 123 AFFSHTSVNMWSTSETQKILFKQSLGPNLNPYLSHLAKLYVANGSGVEVRESISGSGVF 182  
 QY 120 AGKIISCIPPGF-AAQNISIAQATMFPHVIADRVRLPIEVPLEDVNRVLFHNDNAPT 178  
 Db 183 GGLAAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEPVFAIPDLRSNLYHLMSDIDT 239  
 QY 179 MRLVCMY-----TPLRASGSSGTPDFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPSV 234  
 Db 240 TSLVMVYNDLINPYANDTNSGCG---IVT--VETKPGDFKFKLLKPPG----- 284  
 QY 235 PNLPLNTLSNRVPS-LI---KSMVMVRDGHQMVQ-----FQ-NGRVTLGDLQLOG-- 279  
 Db 285 -----SMLTHGCVPSDLIPKSSSLIGNRHWSDITDFIIRPFVFOQANRHFDFNOETAGWS 339  
 QY 280 -----ITPTSASOLCKI-----RGSVFHANGNGYN 305  
 Db 340 TPRFRPITTVSESNNKLGICGATVATYIVPGIPDGPWDTTIPQLTPAGIYSITASNOTD 399

```

RESULT 5
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racanello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

```

RESULT 6  
US-08-461-503-2  
; Sequence 2, Application US/08461503  
; Patent No. 5834302  
; GENERAL INFORMATION:  
; APPLICANT: Racanietello, Vincent

```

; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 36607-D-PCT-US
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 423523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-461-503-2

```

```

Query Match      5.1%; Score 149; DB 2; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

```

```

QY 15 TSGAGOLVPEANTAEPISEMPVAGATAAATAGQV-NM-----IDPWIMNNYVQAQPG-- 66
Db 347 TPGSNQYLTSNDHNSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ----EFTISPNNTPGDILFDLQGLPHLNPFSLH-----LAQMYNGWGNMKVYLLAGNAF 118
Db 406 DMRYVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLYNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIIPGFAQONISIAQATMFPVIAADVRLPIEVPLEDVRNVLPH--NNDNA 176
Db 466 ATGKILVAYAPP-QOPPTSREKAMLGTHVIWDLGLQSSCTMVVPWISNVYRTQDTSF 524
QY 177 PTMRLVCMLY-----TPLRASGSSGTDPFVIAGRVLTCPSPDPSFLFL-----V 221
Db 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVNLPL-----LNTLSNRVPSLIKSM 255
Db 577 PQGIEDLTSEVAGALTLSLPKQKQDLPDTKASGPAHSKEVPALTAETGATNPLAPSDT 636
QY 256 VSRDHQWQV-----FQNGRVTLGQLOQTTPTSASQCKIRGSVFH----- 297
Db 637 VQTRH--VVQRRSRSESTIESFFARGACVAIIEDVNEQPTTRAQKLFAMWRITYKDTVOL 694
QY 298 -----ANGNGYNLTDLGSPYHAFESPAPIGFPDLGECDW 333
Db 695 RRKLEFFYSRDEMEFTVVVTANFNANNGHALNQVQIMYIPPGAPTPKSWDDY---TW 751
QY 334 HMEASPTQNTGDVTKQNVKQESAFAPHLGTIOA-----DGLSDVSYNNTNNIAKLG 386
Db 752 QTSNSISIFTYCAAPARISV-----PYVGLANAYSHFYDGFAGVPLKTDANQIG 802

```

```

RESULT 7
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racanliello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release 1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-465-250-2

```

```

Query Match      5.1%; Score 149; DB 4; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

```

```

QY 15 TSGAGOLVPEANTAEPISEMPVAGATAAATAGQV-NM-----IDPWIMNNYVQAQPG-- 66
Db 347 TPGSNQYLTSNDHNSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ----EFTISPNNTPGDILFDLQGLPHLNPFSLH-----LAQMYNGWGNMKVYLLAGNAF 118
Db 406 DMRYVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLYNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIIPGFAQONISIAQATMFPVIAADVRLPIEVPLEDVRNVLPH--NNDNA 176
Db 466 ATGKILVAYAPP-QOPPTSREKAMLGTHVIWDLGLQSSCTMVVPWISNVYRTQDTSF 524
QY 177 PTMRLVCMLY-----TPLRASGSSGTDPFVIAGRVLTCPSPDPSFLFL-----V 221
Db 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVNLPL-----LNTLSNRVPSLIKSM 255
Db 577 PQGIEDLTSEVAGALTLSLPKQKQDLPDTKASGPAHSKEVPALTAETGATNPLAPSDT 636
QY 256 VSRDHQWQV-----FQNGRVTLGQLOQTTPTSASQCKIRGSVFH----- 297
Db 637 VQTRH--VVQRRSRSESTIESFFARGACVAIIEDVNEQPTTRAQKLFAMWRITYKDTVOL 694
QY 298 -----ANGNGYNLTDLGSPYHAFESPAPIGFPDLGECDW 333

```

Db 695 RRKLEFFYSRDEHFEFTVVTANFTHANNGHALNQVQIMYIPPGAPTPKSWDDY---TW 751  
QY 334 HMEASPTQFNTGDIKQINVKQESAFAPHLGTIAQ-----DGLSDVSVNTNMIKLG 386  
Db 752 QTSSNPISIFTYGAAPARISV-----PVGLANAYSHEFDGFAKVPKLTANDQIG 802  
RESULT 8  
US-09-423-890-2  
; Sequence 2, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:  
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPCP  
; CURRENT APPLICATION NUMBER: US/09/423,890  
; CURRENT FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USSN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USSN 60/099,165  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1302  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-423-890-2  
Query Match 3.7%; Score 106; DB 4; Length 1302;  
Best Local Similarity 18.7%; Pred. No. 0.24;  
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;  
QY 1 MMASKDAPTNMDGTSCAGOLVPEANTAEPISEMPVAG-AATAAATAGQVNMIDPIMNN 59  
Db 400 LLLANGESTGNSGGSG-GSL--SAGAAAGSSQPSISGDVVEACCSVLSTVCADP-VYKV 455  
QY 60 YVQA---POGEFTISPNTPGD-ILFDLQGLPHLNPFL-----SHLAQMY 100  
Db 456 YVAALKTLRAMLVYTPCHSLAEIKLQRLRPVVDILVKCADANSRTSLSSTVLEIC 515  
QY 101 NGWGNMKV--KVLGAGNAFTAG-KIIISCIPGFAAQN-----ISIAQA 142  
Db 516 NQAGLAVGREILKAGSIGVGVDYVLSCLGNQAESNMWQELLGRCLLDRLLEFPA 575  
QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNVL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196  
Db 576 EFYPHIVSTDVSOAEPVEIRYKLLSLTLFALQSIDNSHSMWGKLSRIY----- 625  
QY 197 GTPDFVIAGRLVTCPSPDFSFLFVPPNVBQKTKPFVSPNPLPLNTLSNRVPSLIKSMV 256  
Db 626 -----LSSARMVTAVPAVFSKLYTM-----LNASGSTHFTMRRLMA 663  
QY 257 SRDHGMQV-FQNG-RVTLDG---QLQGTTPTSASQLCKIRGSVFHANGNGYNLTELD- 310  
Db 664 IADEVEIAEVIQLGVEDTVGDHDSLQALAPASCLSENSLSLEHTVHREKTKGLSATRLSA 723  
QY 311 -----GSPHAFESPAPIGFPDPLGECDDHWEAS 338  
Db 724 SSEDISDRLAGSVGLPSSTTTQPKPAVOTKGRPHSQCLNSSLPLSHAQL-----MFA 777  
QY 339 PTTQFNTGDIKQINVKQESAFAP 362  
Db 778 PSAPCSSAPSPDLSKRPQAFVP 801  
RESULT 9  
US-09-423-890-8  
; Sequence 8, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Exter  
; FILE REFERENCE: CPI-004DVC3  
; CURRENT APPLICATION NUMBER: US/08/628,829A  
; CURRENT FILING DATE: 1996-04-05  
; EARLIER APPLICATION NUMBER: 08/440,421  
; EARLIER FILING DATE: 1995-05-15  
; EARLIER APPLICATION NUMBER: 08/323,460  
; EARLIER FILING DATE: 1994-10-14  
; EARLIER APPLICATION NUMBER: 08/049,254  
; EARLIER FILING DATE: 1993-05-15  
; EARLIER APPLICATION NUMBER: 08/410,602  
; EARLIER FILING DATE: 1995-04-24

; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPCP  
; CURRENT APPLICATION NUMBER: US/09/423,890  
; CURRENT FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USSN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USSN 60/099,165  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 1493  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-423-890-8  
Query Match 3.7%; Score 106; DB 4; Length 1493;  
Best Local Similarity 18.7%; Pred. No. 0.31;  
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;  
QY 1 MMASKDAPTNMDGTSCAGOLVPEANTAEPISEMPVAG-AATAAATAGQVNMIDPIMNN 59  
Db 591 LLLANGESTGNSGGSG-GSL--SAGAAAGSSQPSISGDVVEACCSVLSTVCADP-VYKV 646  
QY 60 YVQA---POGEFTISPNTPGD-ILFDLQGLPHLNPFL-----SHLAQMY 100  
Db 647 YVAALKTLRAMLVYTPCHSLAEIKLQRLRPVVDILVKCADANSRTSLSSTVLEIC 706  
QY 101 NGWGNMKV--KVLGAGNAFTAG-KIIISCIPGFAAQN-----ISIAQA 142  
Db 707 KQAGLAVGREILKAGSIGVGVDYVLSCLGNQAESNMWQELLGRCLLDRLLEFPA 766  
QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNVL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196  
Db 767 EFYPHIVSTDVSOAEPVEIRYKLLSLTLFALQSIDNSHSMWGKLSRIY----- 816  
QY 197 GTPDFVIAGRLVTCPSPDFSFLFVPPNVBQKTKPFVSPNPLPLNTLSNRVPSLIKSMV 256  
Db 817 -----LSSARMVTAVPAVFSKLYTM-----LNASGSTHFTMRRLMA 854  
QY 257 SRDHGMQV-FQNG-RVTLDG---QLQGTTPTSASQLCKIRGSVFHANGNGYNLTELD- 310  
Db 855 IADEVEIAEVIQLGVEDTVGDHDSLQAVAPTSCLSENSLSLEHTVHREKTKGLSATRLSA 914  
QY 311 -----GSPHAFESPAPIGFPDPLGECDDHWEAS 338  
Db 915 SSEDISDRLAGSVGLPSSTTTQPKPAVOTKGRPHSQCLNSSLPLSHAQL-----MFA 968  
QY 339 PTTQFNTGDIKQINVKQESAFAP 362  
Db 969 PSAPCSSAPSPDLSKRPQAFVP 992  
RESULT 10  
US-08-628-829-4  
; Sequence 4, Application US/08628829A  
; Patent No. 6333170  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Exter  
; FILE REFERENCE: CPI-004DVC3  
; CURRENT APPLICATION NUMBER: US/08/628,829A  
; CURRENT FILING DATE: 1996-04-05  
; EARLIER APPLICATION NUMBER: 08/440,421  
; EARLIER FILING DATE: 1995-05-15  
; EARLIER APPLICATION NUMBER: 08/323,460  
; EARLIER FILING DATE: 1994-10-14  
; EARLIER APPLICATION NUMBER: 08/049,254  
; EARLIER FILING DATE: 1993-05-15  
; EARLIER APPLICATION NUMBER: 08/410,602  
; EARLIER FILING DATE: 1995-04-24

us-09-926-799-1.rai

```

; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-628-829-4

Query Match          3.7%; Score 106; DB 4; Length 1593;
Best Local Similarity 18.7%; Pred. No. 0.34;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20:

Qy 1 MMWASKDAPNTMDGTSGAGOLVPEANTAEPISMPEVAG-AATAAATAGOVNMIDPWIMNN 59
Db 691 LLLANGESTNSGGSG-GSL--SAGASGSSQPSISGDVVEACCSVLIVCADP-VYKV 746
Qy 60 YVQA---PQGEFTISPNTPGD-ILFDLQLGPHLNPFL-----SHLAQMY 100
Db 747 YVAALKTLRAMLVYTPCHSLAERIKLQRLRPVVDTILVKCADANSRTSOLISISTVLELC 806
Qy 143 TMFPHVIA-DVRVLEPIEVLDPVRYNL---FHNNDNAPT---RLVCMLYTPLRASCSS 196
Db 807 KGQAGELAVGREILKAGSIGVGGVYVLSCLIGNQAESNNWQELLGLCLIDRLLEFFA 866
Qy 143 EFYHIVSTDVSOAEPEIRYKLLSLTLTALQSIDNSHSMVGKLSRIY-----916
Db 867 EFYHIVSTDVSOAEPEIRYKLLSLTLTALQSIDNSHSMVGKLSRIY-----916
Qy 197 GTDFEVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVPLNPLNTLSNRVPSLIKSMV 256
Db 917 -----LSSARMTAVPAVFSKLVTM-----LNASGTHFTMRRLMA 954
Qy 257 SRDHGMQV-FQNG-RVTLDG---QLQGTPTTSASQLCKIRGVSFHANGNGYNLTELD- 310
Db 955 IADEVEIAEVQLGVEDTVGHQDSLQNAVPTSCLENSLSLEHTVHREKTKGLSATRLSA 1014
Qy 311 -----GSPYHAFESPAPIGFPDGLGCDWHMEAS 338
Db 1015 SSEDISDRLAGVSLPSTTTEQPKPAVQTKGRPHSQCLNSPLSHAQL-----MPPA 1068
Qy 339 PTTQFNTGDVTKQINVKOESAFAP 362
Db 1069 PSAPCSAPSVPDISKRRQAFVP 1092

RESULT 11
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-09-091-219-24

Query Match          3.7%; Score 106; DB 4; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.64;

Matches 63; Conservative 31; Mismatches 113; Indels 76; Gaps 13:

Qy 5 SKDAPTNMDGTSGAGOLVPEANTAEPISMPE-----PVAGATAAATAGOVNMIDPWIMNNY 60
Db 270 SKLASSAFSGLFGA--LLADKTEETLLDRLTLTRNGHTTSTTOSVGV-----TY 320
Qy 61 VOAPQGEFTISPNTPGDI-----LFDLQLGP-----HL-----NPF 92
Db 321 GYATAEDFVSGPNTSGLETRVVOAERFFKTHLFDWYTSDFSGRCHLLELPTDCHKGVYGS 380
Qy 93 LSHLAQMYNGWGNMKYKVLVLLAGNAFTAGKIIISCIPGFAAQNISIAQATMPPHVIADY 152
Db 381 TDSYAYMRNGW---DVEVTAVGNQFNGCGLLVAMVPGLYSIQKRELYQLTLEPHQFINP 436
Qy 153 R--VLEPIEVLDPVRYN-----LFDLQLGP-----LFDLQLGP-----NPF 92
Db 437 RTNMTAHTITVDFGVGNRYDYQYVHKPWTLLVVMVAVPLTVNTEGAPQIKVYANLPIVHV 496
Qy 192 SG---SSSGCTDPFVIA---CRVLTCP---SPDFSFLFLVPPN 224
Db 497 AGEFPSKEGIEFPVACSDGCGYGLVTTDPKTADPVYGVKVENPPRN 539

RESULT 12
US-09-749-588-2
; Sequence 2, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-749-588-2

Query Match          3.6%; Score 103.5; DB 4; Length 1170;
Best Local Similarity 20.9%; Pred. No. 0.37;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33:

Qy 9 PTNMDGTSGAGOLVPEANTAEPISMPEVAGAAAT-AAATAGOVNMIDPWIMNNYVOAPOQE 67
Db 633 PVRMD---NAVPIVQPPAAQPLQIQ--SGVLTQGSCTPLMVATLHPQVA-----677
Qy 68 FTISPNTPGDIILFDLQLGPHLNPFLSHLAQMYNGWGNMKYKVLVLLAGNAFTAGKIIISC 127
Db 678 -TITPQVA---VPFTLSCAAGRPAALVEQTAALVLAQMPGGTQ-QILLPS---TWOQL----725
Qy 128 IPPGFAAQNTSIAQATMPPHVIADRVVLEPIEVLDPEDVRNVLFNNDNAPTMRVLCMLYT 187
Db 726 --PCVALHN-SVQPTAMPEAMSGO-----QLADWRNAHSHGNQYSTIMQQOSSLTN 775
Qy 188 -----PLRAS-----GSSSGTDPPEVIAGRV--LTCPSP 213
Db 776 HVTLATAQPLNVGVAHVVRQOQSSSLPSKKNKOSAPVSSKSLD--VLPQSVYSLVSSSP 833
Qy 214 ---DFSEFLFLVPPNVEOKTKPFSVPLNPLNTLS-----NSRVPSLIKSMVYS 257
Db 834 LRTTSSYNSLVP--VQDQHQPITIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 889
Qy 258 RDHGMVQFQNGRVTLQGLQGTPTTSASOLCKIRGSVFHANGN---GYNLTELDGSPY 314
Db 890 RSN--VISYTVNDSPDSLSLSPYSTDTLSALRG-----NSGSLVEGGRVAVDGTGT 942
Qy 315 HAFESPAPIGFPDGLGCDWHMEASPTTQFNTGDVTKQINVKOESAFAPHLGCTIQADGLS- 373
```

Db 943 RTIIVP-PLK-TQLGDCVATQASGLLSNKKPKVPASVSGSSGCCITPTGYRAQRGGTSA 1000  
QY 374 -----DVSNTNMIAKLGWSPVSDGHRGDVDPWVPIRYGSTLTFEAAQLAPPIYPGFGGEA 429  
Db 1001 AQLNLSONQOOSA-----APTQSRSSNPAP-----RRQAFVAP-----LSQA 1040  
QY 430 IVFMSDFPTAHGTNG-----LSVPCTTPOEFVTHFVNEQAPTRGEEA-----ALLHYLDPD- 480  
Db 1041 PYTFQGSPL-HSTGHPHLAPAPAHLPQ--AHLYTYAAPTSAALGSTSSIAHLFSPQG 1097  
QY 481 THNLGEFKLYPEGFMTCVPNSSGTGPTQL 510  
Db 1098 SSRHAAAYTHPSTLVHQVPS--VGPLL 1125  
RESULT 13  
US-08-460-860-4  
; Sequence 4, Application US/08460860  
; Patent No. 5665584  
; GENERAL INFORMATION:  
; APPLICANT: WATAMOTO, OSAMU  
; APPLICANT: WATARAI, TERUO  
; APPLICANT: MIZUSAWA, KIYOSHI  
; APPLICANT: NAKANO, EIICHI  
; TITLE OF INVENTION: A DNA FRAGMENT CONTAINING A TANNASE  
; TITLE OF INVENTION: GENE, A RECOMBINANT PLASMID, A PROCESS FOR PRODUCING  
; TITLE OF INVENTION: TANNASE, AND A PROMOTER  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,860  
; FILING DATE: 05-JUN-1995  
; -CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: JP 159973/1994  
; FILING DATE: 12-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 83973/1995  
; FILING DATE: 10-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F  
; REFERENCE/DOCKET NUMBER: 7127-001-0X  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 588 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-460-860-4  
Query Match 3.5%; Score 102.5; DB 1; Length 588;  
Best Local Similarity 21.7%; Pred. No. 0.15;  
Matches 66; Conservative 41; Mismatches 100; Indels 97; Gaps 18;  
QY 273 LDQLOGTTPTTSASQCKIRGVSFHANGNGYNLTGDSPPYH-AFESPAPIGFPDLGEC 331  
Db 264 LDGRTDGV--VSRDLDCKL-----NFNLTSIIIGEPYCAAGTSTSLGF----- 304

QY 332 DHMEASPTTFQNTGVIKQINVKQE-----SAFAP-HLCTIQADGLSDYSV-----N 378  
Db 305 -----GFSNG---KRSNVKROAEGSTTSYQPAQNGCTVARGVAVAAIYDGLHNS 351  
QY 379 TNMIAKLGW--VSPVSDG---HRGDVDPVW--IPRYGS-----TLTEAAQLAPPI 421  
Db 352 KGERAYLSWQIASELSDAETEYNSDTGKNELNIPSTCGEYVTKFIQLLNJNLSLNNVT 411  
QY 422 YPGFGEAIVFEFMSDFPIAHGTNGLSVPCCTIPO-----EFVTHFVNEQAPTRGGAAL 473  
Db 412 Y-----DTLVDMWNTGNVRYMD---SLQTTLPDLTFQSSGGKLLHYHGSDPSIPAASS 463  
QY 474 LHYLDP-----DTHRNLFGEFKLYPEGFMTCVPNSSGTGPTQLPIGVYFV 518  
Db 464 VHYVQAVRSVMYGDKTEEEALEADWYQFYLLIP-GAAHCGTNSLQPGP--YPENNMEIM 520  
QY 519 VSWV 522  
Db 521 IDWV 524  
RESULT 14  
US-09-171-699-10  
; Sequence 10, Application US/09171699  
; Patent No. 6448389  
; GENERAL INFORMATION:  
; APPLICANT: The Wistar Institute of, Anatomy & Biology  
; Gonczol, Eva  
; Berencsi, Klara  
; Kari, Csaba  
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and  
; Uses Therefor  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Center, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171,699  
; FILING DATE: 19-Jan-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/015,717  
; FILING DATE: 23-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kodroff, Cathy A.  
; REGISTRATION NUMBER: 33,980  
; REFERENCE/DOCKET NUMBER: WST66APCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1048 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-171-699-10  
Query Match 3.5%; Score 102; DB 4; Length 1048;  
Best Local Similarity 21.8%; Pred. No. 0.44;  
Matches 120; Conservative 50; Mismatches 214; Indels 166; Gaps 25;





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:56:00 ; Search time 12 Seconds  
(without alignments)  
902.816 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDPTNMDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	106	3.7	1302	12	US-10-000-864-2
2	106	3.7	1493	10	US-09-858-754-3
3	106	3.7	1493	12	US-10-000-864-8
4	103.5	3.6	1170	12	US-10-135-687-2
5	103.5	3.6	1210	10	US-09-860-352A-2
6	100.5	3.5	1070	10	US-09-735-367B-6
7	100	3.5	828	9	US-09-738-626-5038
8	100	3.5	1601	10	US-09-862-027-40
9	100	3.5	2005	10	US-09-735-367B-3
10	100	3.5	2063	10	US-09-735-367B-2
11	98	3.4	1209	12	US-10-135-687-4
12	98	3.4	1493	10	US-09-858-754-4
13	97	3.3	956	9	US-10-121-032-63
14	96	3.3	1344	9	US-09-738-626-6888
15	95.5	3.3	678	9	US-09-712-363-158
16	95	3.3	504	9	US-09-738-626-3680
17	94.5	3.3	662	10	US-09-924-358-11
18	94	3.2	1477	9	US-10-092-880-4
19	93.5	3.2	498	10	US-09-925-297-673

20	93	3.2	581	9	US-09-738-626-5327	Sequence 5327, Ap
21	92.5	3.2	477	10	US-09-972-715-9	Sequence 9, Appli
22	92.5	3.2	594	10	US-09-801-368-112	Sequence 112, App
23	92.5	3.2	595	10	US-09-768-826-48	Sequence 48, Appl
24	92.5	3.2	2308	12	US-10-000-934-2	Sequence 2, Appli
25	92	3.2	996	8	US-08-910-386A-5	Sequence 5, Appli
26	91.5	3.2	157	10	US-09-893-737-18	Sequence 18, Appl
27	91.5	3.2	910	9	US-09-855-754-5	Sequence 5, Appli
28	91.5	3.2	3472	9	US-10-027-806-4	Sequence 4, Appli
29	91.5	3.2	3472	9	US-10-034-623-4	Sequence 4, Appli
30	91.5	3.2	5795	10	US-09-815-242-12610	Sequence 12610, A
31	91	3.1	1263	10	US-09-864-761-46846	Sequence 46846, A
32	91	3.1	1706	10	US-09-864-761-46862	Sequence 46862, A
33	91	3.1	4563	9	US-09-870-759-128	Sequence 128, App
34	90.5	3.1	535	10	US-09-768-826-38	Sequence 38, Appl
35	90	3.1	505	1	US-08-976-063C-22	Sequence 22, Appl
36	90	3.1	925	10	US-09-452-380-4	Sequence 4, Appli
37	90	3.1	936	10	US-09-452-380-3	Sequence 3, Appli
38	90	3.1	1298	9	US-09-825-288A-2	Sequence 2, Appli
39	89.5	3.1	400	9	US-10-209-264-3	Sequence 3, Appli
40	89.5	3.1	542	10	US-09-871-212-9	Sequence 9, Appli
41	89.5	3.1	578	9	US-09-891-139A-3	Sequence 3, Appli
42	89.5	3.1	6304	9	US-10-147-026-16	Sequence 16, Appl
43	89	3.1	1781	10	US-09-738-877-3	Sequence 3, Appli
44	88	3.0	400	10	US-09-073-009-126	Sequence 126, App
45	88	3.0	400	10	US-09-793-306-126	Sequence 126, App

ALIGNMENTS

RESULT 1  
US-10-000-864-2  
; Sequence 2, Application US/10000864  
; Patent No. US20020146798A1  
; GENERAL INFORMATION:  
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPC  
; CURRENT APPLICATION NUMBER: US/10/000,864  
; EARLIER FILING DATE: 2001-10-31  
; EARLIER FILING DATE: 2000-06-03  
; EARLIER FILING DATE: 2000-06-03  
; EARLIER FILING DATE: 1999-03-15  
; EARLIER APPLICATION NUMBER: USSN 60/078,153  
; EARLIER FILING DATE: 1998-03-16  
; EARLIER APPLICATION NUMBER: USSN 60/099,165  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1302  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-864-2

Query Match 3.7%; Score 106; DB 12; Length 1302;  
Best Local Similarity 18.7%; Pred. No. 0.79;  
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

Qy	1	MMASKDPTNMDGTSGAGOLVPEANTAPISMEPVAG-AATAAATAGVNMIDPMINN	59
Db	400	LLUANGESTGSGGGG-GSL--SAGAAGSGQSPISGDVVEACCSVLSIVCADP-VYKV	455
Qy	60	YVQA----POGEETISPNPTPGD-ILFDLQLGHLNPLF-----SHLAQMY	100
Db	456	YVAALKTRLAMLYVTPCHSLAEIRIKLQRLRPVVDITLVKCADANSRTSLSSTVLELC	515
Qy	101	NGWVGNNKV--KVLLAGNAFTAG-KTIICIPPGFAQN-----ISIAQA	142
Db	516	NGQAGKLAVGREILKAGSIGVGVGVVLSILCNQAESNNWQELLGLRLCLIDRLLEFPA	575

```

QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196
Db 576 EYFPHIVSTDVSOAEVEIRYKLLSLTLFALQSIDNSHSMVGKLSRIY----- 625
QY 197 GTDPFVIAGRVLTCPSDFSLFLVPPNVNQKTKPFSVPNPLPLNTLSNRVPSLIKSMV 256
Db 626 -----LSSARMVTAVPAVESKLVTM-----LNAGSGSTHFTMRRLMA 663
QY 257 SRDHGMQV-FQNG-RVTLDG---QLQGTTPTSASOLCKIRGSVFHANGNGYNLTELD- 310
Db 664 IADEVEIAEVIQLGVEDTVGHODSLQALAPASCLNSLSLEHTVHREKTGKLSATRLSA 723
QY 311 -----GSPHAFESPAPICFPDGLGCDWHMEAS 338
Db 724 SSEDISDRLAGVSVGLPSSTTTTQPKPAVQTKGRPHSQCLNSSPLSHAOL-----MFA 777

QY 339 PTTQFNTGDVIKOINVKQESAFAP 362
Db 778 PSAPCSSAPSVDPDISKHRPQAFVP 801

RESULT 2
US-09-858-754-3
; Sequence 3, Application US/09858754
; Patent No. US2002005130A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING APOPTOSIS
; FILE REFERENCE: CPI-042
; CURRENT APPLICATION NUMBER: US/09/858,754
; PRIOR FILING DATE: 2001-05-16
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 3
; LENGTH: 1493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-754-3

Query Match 3.7%; Score 106; DB 10; Length 1493;
Best Local Similarity 18.7%; Pred. No. 0.96; Indels 124; Caps 20;
Matches 83; Conservative 72; Mismatches 165;

QY 1 MMWASKDAPTNDGTSGAGQLVPEANTAEPISEMPEVAG-AATAAATAGQVNMIDPWIMNN 59
Db 591 LLLANGESTGSGGSG-GSL--SAGAAGSGSQPSISGDVVEACCSVLISVCADP-VYKV 646
QY 60 YVOA---PQGETTISPNTPGD-ILFDLQGLHLPFL-----SHLAQMY 100
Db 647 YVAALKTLRALMVLVTPCHSLAERIKLRLRPVVDTLVKCADANSRTSOLSIIVLELC 706
QY 101 NGWVGNMKV--KVLGNAFTAG-KIISCIPPGFAAQN-----ISIAQA 142
Db 707 KGAGELAVGREILKAGSIGVGVDYVLSCLIGNQAESNNWQELLGRCLIDRLLEPPA 766
QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196
Db 767 EYFPHIVSTDVSOAEVEIRYKLLSLTLFALQSIDNSHSMVGKLSRIY----- 816
QY 197 GTDPFVIAGRVLTCPSDFSLFLVPPNVNQKTKPFSVPNPLPLNTLSNRVPSLIKSMV 256
Db 817 -----LSSARMVTAVPAVESKLVTM-----LNAGSGSTHFTMRRLMA 854
QY 257 SRDHGMQV-FQNG-RVTLDG---QLQGTTPTSASOLCKIRGSVFHANGNGYNLTELD- 310
Db 855 IADEVEIAEVIQLGVEDTVGHODSLQAVAPTSCLENSLSLEHTVHREKTGKLSATRLSA 914
QY 311 -----GSPHAFESPAPICFPDGLGCDWHMEAS 338

QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNL---FHNNDNAPTM--RLVCMLYTPLRASGSS 196
Db 576 EYFPHIVSTDVSOAEVEIRYKLLSLTLFALQSIDNSHSMVGKLSRIY----- 625
QY 197 GTDPFVIAGRVLTCPSDFSLFLVPPNVNQKTKPFSVPNPLPLNTLSNRVPSLIKSMV 256
Db 626 -----LSSARMVTAVPAVESKLVTM-----LNAGSGSTHFTMRRLMA 663
QY 257 SRDHGMQV-FQNG-RVTLDG---QLQGTTPTSASOLCKIRGSVFHANGNGYNLTELD- 310
Db 664 IADEVEIAEVIQLGVEDTVGHODSLQALAPASCLNSLSLEHTVHREKTGKLSATRLSA 723
QY 311 -----GSPHAFESPAPICFPDGLGCDWHMEAS 338
Db 724 SSEDISDRLAGVSVGLPSSTTTTQPKPAVQTKGRPHSQCLNSSPLSHAOL-----MFA 777

QY 339 PTTQFNTGDVIKOINVKQESAFAP 362
Db 778 PSAPCSSAPSVDPDISKHRPQAFVP 801

RESULT 4
```

```
US-10-135-687-2
; Sequence 2, Application US/10135687
; Patent No. US20020123120A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001068DIV
; CURRENT APPLICATION NUMBER: US/10/135,687
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 09/749,588
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-687-2

Query Match      3.6%; Score 103.5; DB 12; Length 1170;
Best Local Similarity 20.9%; Pred. No. 1.1;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

QY 9 PTNMDCTGAGQLVPEANTAEPISEMPVAGAAT-AAATAGQVNMIDPWIMNNYVQAPOGE 67
Db 633 PVRMD---NAVPIVQAPAAQPLQIQ--SGVLTQGGCTPLMWATLHPQVA----- 677

QY 68 FTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIIISC 127
Db 678 -TITPOYA---VPFTLSCAAGRPALEQTAALVLAQWPGGTQ-QILLPS---TWQOL---- 725

QY 128 IPGFAAQNISTAAQTMFPHVADRVLEPIEVPLEDVNRVLFHNNDNAPTMRVLCMLYT 187
Db 726 --PGVALHN-SVQPTAMIEPMGSGQ-----QLADWRNAHSHGNQYSTIMOQPSLLTN 775

QY 188 -----PLRAS-----GSSSGTDPFVIAGRV--LTCPS 213
Db 776 HVTATAQPLNVGVAVHVRQOQSSSLPSKKKQSAVPSSKSLD--VLPSQVYSLVGS 833

QY 214 ---DFSFLFPLVPNEOKTKPFSVNLPLNTLS-----NSRVPSLIKSMWVS 257
Db -834 LRTTSYNSLVP--VDQHQPIIIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 889

QY 258 RHGQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANGN---GYNLTELDGSPY 314
Db 890 RSN--VISYVTYVNDSPDSSLSSTYSTDLSALRG-----NSGVLEGPGRVVADGTGT 942

QY 315 HAFESAPIGFDPDLGECDDHMEASPTQFNTGDIKQINVKQESAFAPHLGTTIQADGLS- 373
Db 943 RTIIVP-PLK-TQLGCTVATQASGLLSNKTKPVASVSGSSGCCITPTGYRAQRGTS 1000

QY 374 ----DYSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPPIYPGFG 429
Db 1001 AOPNLNSQNOQSSA-----APTSQERSNNAP-----RRQAFVAP-----LSQA 1040

QY 430 IVFFMSDFPIAHGTNG---LSVPCTIPQEFVTHFVNEQAPTRGEA-----ALLHYLDPD- 480
Db 1041 PYTFQHGSP-LHSTGHPLAPAPAHLPQSO--AHLTYAAPTSAALGSTSTIAHLFSPOG 1097

QY 481 THRNLEGEKLYPEGMTCVPSNSSGTGCPQTL 510
Db 1098 SSRHAAAYTHPSTLVHQPVVS--VGPSLL 1125

RESULT 5
US-09-860-352A-2
; Sequence 2, Application US/09860352A
; Patent No. US20020132785A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Curtis, Rory
```

```
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: 13305 NOVEL PROTEIN KINASE MOLECULES AND
; FILE REFERENCE: 38155-20016.00
; CURRENT APPLICATION NUMBER: US/09/860,352A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,301
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-352A-2

Query Match      3.6%; Score 103.5; DB 10; Length 1210;
Best Local Similarity 20.9%; Pred. No. 1.2;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

QY 9 PTNMDCTGAGQLVPEANTAEPISEMPVAGAAT-AAATAGQVNMIDPWIMNNYVQAPOGE 67
Db 673 PVRMD---NAVPIVQAPAAQPLQIQ--SGVLTQGGCTPLMWATLHPQVA----- 717

QY 68 FTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIIISC 127
Db 718 -TITPOYA---VPFTLSCAAGRPALEQTAALVLAQWPGGTQ-QILLPS---TWQOL---- 765

QY 128 IPGFAAQNISTAAQTMFPHVADRVLEPIEVPLEDVNRVLFHNNDNAPTMRVLCMLYT 187
Db 766 --PGVALHN-SVQPTAMIEPMGSGQ-----QLADWRNAHSHGNQYSTIMOQPSLLTN 815

QY 188 -----PLRAS-----GSSSGTDPFVIAGRV--LTCPS 213
Db 816 HVTATAQPLNVGVAVHVRQOQSSSLPSKKKQSAVPSSKSLD--VLPSQVYSLVGS 873

QY 214 ---DFSFLFPLVPNEOKTKPFSVNLPLNTLS-----NSRVPSLIKSMWVS 257
Db 874 LRTTSYNSLVP--VDQHQPIIIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 929

QY 258 RHGQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANGN---GYNLTELDGSPY 314
Db 930 RSN--VISYVTYVNDSPDSSLSSTYSTDLSALRG-----NSGVLEGPGRVVADGTGT 982

QY 315 HAFESAPIGFDPDLGECDDHMEASPTQFNTGDIKQINVKQESAFAPHLGTTIQADGLS- 373
Db 983 RTIIVP-PLK-TQLGCTVATQASGLLSNKTKPVASVSGSSGCCITPTGYRAQRGTS 1040

QY 374 ----DYSVNTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQLAPPIYPGFG 429
Db 1041 AOPNLNSQNOQSSA-----APTSQERSNNAP-----RRQAFVAP-----LSQA 1080

QY 430 IVFFMSDFPIAHGTNG---LSVPCTIPQEFVTHFVNEQAPTRGEA-----ALLHYLDPD- 480
Db 1081 PYTFQHGSP-LHSTGHPLAPAPAHLPQSO--AHLTYAAPTSAALGSTSTIAHLFSPOG 1137

QY 481 THRNLEGEKLYPEGMTCVPSNSSGTGCPQTL 510
Db 1138 SSRHAAAYTHPSTLVHQPVVS--VGPSLL 1165

RESULT 6
US-09-735-367B-6
; Sequence 6, Application US/09735367B
; Patent No. US20020151477A1
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Jan-Ake
; APPLICANT: Calica, Francoise
; APPLICANT: Antonsson, Per
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR
; FILE REFERENCE: 102093-100
; CURRENT APPLICATION NUMBER: US/09/735,367B
; CURRENT FILING DATE: 2000-12-12
```

```
; PRIOR APPLICATION NUMBER: US 60/174,544
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1070
; TYPE: PR
; ORGANISM: mammal
US-09-735-367B-6

Query Match          3.5%; Score 100.5; DB 10; Length 1070;
Best Local Similarity 18.8%; Pred. No. 1.8;
Matches 102; Conservative 56; Mismatches 188; Indels 197; Gaps 25;

QY 14 GTGAGQLVPEANTARPISENPVAGATAAATAGQV-----NMIDPW-----MNNY 60
Db 542 GNSGAPQLQANQVQH-----AGGAGGPPQONQOVSHGPPNMQPMSLGMTHGNMNNQ 594
QY 61 VQAPQGEFTISPNTPGDILDLQLGP-----HLNPFSLHAQMYNGWGNMKVKV 111
Db 595 QAGTSGVQVNLNMOG-----PQOQPPSOLMGMHQOIIVPSQGMVQO---QGTINPONPMI 649
QY 112 LLAGNAFTAGKIIISCIPPGFAAQNISIAQATMFPHVIAADVRLVLEPIEVPLEDRVNLVLFH 171
Db 650 LSRQAQLMPQGMVNV--PP---SQNLGPPSPQM-----677
QY 172 NDNAPTMRLLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQ-KTK 230
Db 678 -----TPPKQMLSQQGPQMMAHPNQM-----GPOQVLLQONPMIEQIMTN 719
QY 231 PFSVNPILPLNTLSNRV---PSLIKSMVSRDHGMVQFQNGRVTLDGQL---QGTTPTS 284
Db 720 QMOGNKQOFTNQOSNYMPGPAQIMRGPTNMOGNVQVFTG---QMSQMLPQGGPVNNS 776
QY 285 ASQCKIRGVSFVHANGGNGYNLTLDGSPYHAFESPAPIGPFDLGECDWHM---EASPTT 341
Db 777 PSQVWGIQGVLRPPG-----PSP-----HMAQQHGDPAT 806
QY 342 QFNT-----GVVVKOINVKQESAFAPHLGTTQADGLS-----373
Db 807 TANNDVLSLQMMQDVSIQOTNMPVPHVQAMQGNASGNSHFSGHGMSNAPFSGAPNGQM 866
QY 374 -----DVSVNTNMLAKLGWSPVSDGHRGVDVWVPIPRYGSTLTERAQL-----417
Db 867 SCGQNPQFPVKNKDVLTSLPLVNL-LQSDISAGHFG-----VNNKQNTNANKPKKKKPP 920
QY 418 -----APTYPPGGEALVFFMSDPPIAHGTNGLSLVPCPTIPEQFVTHFVNE-QA 465
Db 921 RKKKNSQDNLNTPDTRPAGLEEA-----DQPLPGEGQISLDNSGPK--LPEFSNRPPA 972
QY 466 PTR 468
Db 973 PSQ 975

RESULT 7
US-09-738-626-5038
; Sequence 5038, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OKHAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
```

```
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: Jp 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: Jp 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: Jp 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5038
; LENGTH: 828
; TYPE: PR
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5038

Query Match          3.5%; Score 100; DB 9; Length 828;
Best Local Similarity 22.4%; Pred. No. 1.4;
Matches 106; Conservative 50; Mismatches 143; Indels 174; Gaps 28;

QY 111 VLLAGNAFTAGKIIISCIPPGFAAQNISIAQATMFPHVIAADVRLVLEPIEVPLEDRVNL 169
Db 394 VEIAGGTVDAGRTLVDGP---AMQPTMKVTR--PSELACVDYSAETVIARLEEVGTV 448
QY 170 FHIND-----NAPTRM-----LVCMLYTPLRASGSSSGTDPFVI---AGRVLT-----209
Db 449 AVSGDTLEVTPPTWRGDLTMSADLVEEVLRLEGLEA---IPTIITAPAGRGLTDAQKRRR 506
QY 210 -----CPSPDFSFLFLVPPNV-----EOKTRPFSVNPILPL-----NTL 242
Db 507 AVCHALAVAGVAEIITPSP-----FMDPEFVDVWGCLAADDERRTKTVSVLN-PLEARNVL 559
QY 243 SNRSVPSLIKSM--VSRDHG-----QMVQFQNGRVTLDGLOGTTP TSA-----S 286
Db 560 STSLLLPSMLDAVKRNVARCHNDFSLFGLQOVAFEHG-----SGVSPMPSVASRPEES 611
QY 287 QLCCKIRGVSFVHANGGNGYNLTLDGSPYHA-----FESPAPIG---FPDLGECDW 333
Db 612 VVAEL-----VDSLNPQPLHVATVGTGNTIEFEGPWGKGRAYTFADA-----652
QY 334 HMEASPTTQFNTQDVVVKOINVKQESAFAPHLGTTQA---DGLSDVSVVNTNMLAKLGWSP 390
Db 653 -IESARAVARAAGVTLELANA---DALPWHGRCALLIDG-----TP 691
QY 391 VSDGHRGVDVWVPIPRYG-----STLTERAQLAPPIYPPGGEALVFFMSDP 438
Db 692 V--GYAGELHPQILERAGLPARTCAMELDSLALPLVENLPAPV-----LSSEP 737
QY 439 IAHGTNGLSVPCPTIPEQFVTHFVNEQAPTRGEAALLHYLDPPDTHRN--LGEPK 489
Db 738 ALHQDIALVDEIIPAEVDYRAVVEAGAGELIETVELF---DVRSEQLGENK 786

RESULT 8
US-09-862-027-40
; Sequence 40, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1601
; TYPE: PR
; ORGANISM: C. elegans
US-09-862-027-40

Query Match          3.5%; Score 100; DB 10; Length 1601;
```

Qy	172	NNDNAPTRLVCMILYTPLRASGSSSGDPFVIAGRVLTCPSPDFSFLLFLVPPNVQEQ---KTK	230
Db	620	-----TTPKQMLSQOGPQMMAPIHNM-----GPOGOVLLQONPWIEQIMTN	661
Qy	231	PFSVPNPLNTLSNSRV---PSLTKSMVSRDHGMQVQFONGRVTLDGQL---QGTTPPTS	284
Db	662	QMOGNKQOFNTQNSNVMYPGPAQIMRGPTPNMOCNMVQFTG---QMSGOMLPQOGPVNNS	718
Qy	285	ASOLCKIRGSVFHANGNGYNLTLDGSPYHAFESPAPIGFPDGLGECDDWHM---EASPTT	341
Db	719	PSQVYMGIOGOVLRPPG-----PSP-----HMAOOHQDPAT	748
Qy	342	QFNT---GDVIKQINVKQESAFAPHLTGTTQADGLS-----	373
Db	749	TANNDVLSLQMWMPDVSTQOTNMVPPHVOAMQGSASGNSAGHSFGHGMFENAPFSGAPNCNM	808
Qy	374	-----DVSVNTNMIAKLGWSPVSDGHRGDVDPMVTPRYGSLTLEAAQL-----	417
Db	809	SCGQNPFPVKNKDVLTSPLLVNL-LQSDISAGHFG-----VNNKQNTNANKPKKKKPP	862
Qy	418	-----APPLYPPGGEAIVFEMSDFFPAHTNGTSLVPTPIQ	454
Db	863	RKKNSOODLNTPTDRAGLEEA-----DOPPLPGEGISLDSGPK	904
RESULT 10			
US-09-735-367B-2			
; Sequence 2, Application US/09735367B			
; Patent No. US20020151477A1			
; GENERAL INFORMATION:			
; APPLICANT: Gustafsson, Jan-Ake			
; APPLICANT: Calira, Francoise			
; APPLICANT: Antonsson, Per			
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR			
; FILE REFERENCE: 102093-100			
; CURRENT APPLICATION NUMBER: US/09/735,367B			
; CURRENT FILING DATE: 2000-12-12			
; PRIOR APPLICATION NUMBER: US 60/174,544			
; PRIOR FILING DATE: 2000-01-05			
; NUMBER OF SEQ ID NOS: 18			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 2063			
; TYPE: PRT			
; ORGANISM: Human			
US-09-735-367B-2			
Query Match 3.5%; Score 100; DB 10; Length 2063;			
Best Local Similarity 18.6%; Pred. No. 5.3;			
Matches 98; Conservative 53; Mismatches 183; Indels 194; Gaps			
Qy	14	GTSAGAGQLVPEANTAEPISEMPVAGATAAATAGV-----NMIDPWI-----MNNY	60
Db	542	GNSGAPOLQANQNVQH-----AGGACAGPPQNMQOVSHGPPNMMPQSLMGTGHNNNQ	594
Qy	61	VQAPQGEFTTSPNNTPGDILFDLQLGP-----HLPFLSHLAQMYNGWGNKKVKV	111
Db	595	QAGTSGVPQVNLNMQG-----QQQGPSPSLMGMHQOIVPSQGMVQO---QGTLPNQPMI	649
Qy	112	LLAGNAFTAGKIIISCIPPGFAAONISIAOATMPHVIADRVLEPIEVPLEDVRNVLFH	171
Db	650	LSRAQLMPQGMVYN---PP---SQNLGSPQRM-----	677
Qy	172	NNDNAPTRLVCMILYTPLRASGSSSGDPFVIAGRVLTCPSPDFSFLLVPPNVQEQ---KTK	230
Db	678	-----TTPKQMLSQOGPQMMAPIHNM-----GPOGOVLLQONPWIEQIMTN	719
Qy	231	PFSVPNPLNTLSNSRV---PSLTKSMVSRDHGMQVQFONGRVTLDGQL---QGTTPPTS	284
Db	720	QMOGNKQOFNTQNSNVMYPGPAQIMRGPTPNMOCNMVQFTG---QMSGOMLPQOGPVNNS	776
Qy	285	ASOLCKIRGSVFHANGNGYNLTLDGSPYHAFESPAPIGFPDGLGECDDWHM---EASPTT	341

Db 777 PSQVMIQGVLRPPG-----PSP-----HMAQQRGDEAT 806  
Qy 342 QFNT-----GDVTKQINVKQESAFAPHLGTIQADGLS----- 373  
Db 807 TANNDVSLSQMMPDVSIQDTNMVPHVQAMQNSASGNHFSHGMSFNAPPFGAPNGQM 866  
Qy 374 -----DVSNTNNIAKLGNVSPVSDGHRGDVDPWIPRYGSTLTLEAAQL----- 417  
Db 867 SCQNGFGPVNKDVTLTSLVNL-LQSDISAGHEG-----VNNKQNTNANKPKKKPP 920  
Qy 418 -----APPYIPPGFGEAIVFMSDFPIAHGTNGLSVPCITPQ 454  
Db 921 RKKKNSQDLNTPDTRPAGLEEA-----DQPLPGEQGISLDSNGPK 962

RESULT 11  
US-10-135-687-4  
; Sequence 4, Application US/10135687  
; Patent No. US2002012120A1  
; GENERAL INFORMATION:  
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001068DIV  
; CURRENT APPLICATION NUMBER: US/10/135,687  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: 09/749,588  
; PRIOR FILING DATE: 2000-12-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1209  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-135-687-4

Query Match 3.4%; Score 98; DB 12; Length 1209;  
Best Local Similarity 20.9%; Pred. No. 3.6;  
Matches 119; Conservative 67; Mismatches 239; Indels 144; Gaps 32;

Qy 9 PTNMDGTSCAQLYPEANTAEPISEMPVAGAAAT-AAATAGOVNMIDPWIMNMYVQAPQ-- 65  
Db 673 PVRMD---NAVPIVQAPAAQLQIQ--SGVLTOGSCPTLMVATLHQVA---TITQYA 724  
Qy 66 GEFTTISPNTFGDILFLQLGPHLNPFLSHLAQMYNGHVGNNKVKVLLAGNAFTAGKIII 125  
Db 725 VPFTLSCAGRPA-----LVEQTAAVLQAMPGGTQ-QILLP-----SAWQOL- 764  
Qy 126 SCIPPGFAAQNISIAQATMFHVIADVRLPEIEVPLEDVRLFNHNDNAPTMRVLCML 185  
Db 765 ----PGVALHN-SVQPAAVIPEAMGSSQ-----QLADWRNNAISHGNOYSTIMQOQSLL 812  
Qy 186 YT-----PL-----RAGSSS-----GTDPFVIAGRVLTCPSPDFSF- 218  
Db 813 THNVTLATAQLNVGVVHVRRQQSSSLPSKKNKQSAVSSKSSLEVLPSQVYSLVGSSP 872  
Qy 219 -----FLVPPNVQKTKPFSVNLPLNTLS-----NSRVPSLIKSMVVS 257  
Db 873 LRTTSSYNLSVP--VODQHPIIIPDTSPPVSVTIIKSDTDEEDNKYEPI--SSSLKA 928  
Qy 258 ROHGQWQVQNGRVTLGDLGQGTTPTSASOLCKIRGSVFHANGNGYNLTDLGSPYHAF 317  
Db 929 RSN--VISVTYNDSPDSSLSLSPHSTDTLSALR-----GNSGTLLGPGRPAADG 978  
Qy 318 ESPAPIGFP-----DLGECMDHMEASPTTQFNTGDKVTKQINVKQESAFAPHLGTIQADGLS 373  
Db 979 IGRRTIIVPPLKTLQDCTVATQASGLLSKTKPVASVSGSSGCCITPTGYRAQRGGAS 1038  
Qy 374 DV---SVNTNMIAKLGNVSPVSDGHRGDVDPWIPRYGSTLTLEAAQAPPIYPGFGAI 430  
Db 1039 AVQPLNLSQNOQS-----SSASTSQBERSNP--APR-----RQQAFAVP-----LSQAP 1080

Qy 431 VFMSDFPIAHGTNG---LSVPCTIPIQEFVTHFVNEQAPTRGEA-----ALLHYLDPD-T 481  
Db 1081 YAFQHGSP-L-HSTGHPLAPAPAHLPSPQ---PHLYTYAAPTSAALGSTSIAHLFFPQGS 1137  
Qy 482 HRNLGEFKLYPEGEMTCVPNSSGTGPOTL 510  
Db 1138 SRHAAAYTHPSTILVHQVPVS--VGPSLL 1164

RESULT 12  
US-09-858-754-4  
; Sequence 4, Application US/09858754  
; Patent No. US20020055130A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING APOPTOSIS  
; FILE REFERENCE: CPI-042  
; CURRENT APPLICATION NUMBER: US/09/858,754  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/023,130  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: 60/039,740  
; PRIOR FILING DATE: 1997-02-14  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1493  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-858-754-4

Query Match 3.4%; Score 98; DB 10; Length 1493;  
Best Local Similarity 18.2%; Pred. No. 4.9;  
Matches 83; Conservative 58; Mismatches 129; Indels 186; Gaps 22;

Qy 93 LSHLAQMYNGWGNMKV--KVLLAGNAFTAG-KIISCIPPGFAAQN----- 136  
Db 694 ISTLLELCKGQAGELAVGREILKAGSIGVGVYVLSILGNQAESNNWQELLGRCLUID 753  
Qy 137 --ISIAQATMFHVIA-DVRVLEPIEVLPLEDVRLNVL---FHNNDNAPT--RLVCMLYTP 188  
Db 754 RLLETLSAEFYPHIVSTDVDSQAEPIEIRKLLSLLAFALQSIDNSHSMVGKLSRRIV-- 811  
Qy 189 LRAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVNLPLNTLSNRVP 248  
Db 812 -----LSSARWVTVVPLFSKLVTM-----LSAGSSSHPA 841  
Qy 249 SLIKSMVSRDH---GQWQFQNGRVTLDSQ---LQGTTPTSASOLCKIRGSVFHANGN 302  
Db 842 RMRRLMATADEVEIAEVIQL-GSEDTLGGQDSSOALAPPYPPESSSLEHTAHVEKTK 900  
Qy 303 GYNLTLDGSPYHAFSPA--PIGFPDLGECMDHMEASPTTQFNTGDKVTKQINVKQESAF 360  
Db 901 GLKATRLSASSEDISDRLAGVSVGLP-----SSATTE----- 932  
Qy 361 APHLGTTQADGLSDSVSVNTNMIAKLGNVSPVSDGHRGDVDPWIPRYGSTLTLEAAQLAPP 420  
Db 933 -QPKPTVQTKG-----RPHSQCLNSSPLSP 957  
Qy 421 --LYPPGFGAIVFFMSDFPIAHGTNGLSVPC-----TIQEFVTHFVNEQAPTRGEAALL 474  
Db 958 QLMPF-----AISAPCSSAFSVFAGSVT----- 980  
Qy 475 HYLPDPTHRLNGEFLKYPEGFMTC-VPNSSGTGPQT 509  
Db 981 ---DASKHR-----PRAFVPCIKIPAS---PQT 1002

RESULT 13  
US-10-121-032-63  
; Sequence 63, Application US/10121032  
; Patent No. US2002015550A1

GENERAL INFORMATION:  
APPLICANT: Bylina, Edward J.  
TITLE OF INVENTION: GLYCOSIDASE ENZYMES  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
STREET: 4365 Executive Drive, Suite 1600  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/121,032  
FILING DATE: 09-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/09/134,078  
FILING DATE: 13-AUG-1998  
APPLICATION NUMBER: 08/949,026  
FILING DATE: 10-OCT-1997  
APPLICATION NUMBER: 60/056,916  
FILING DATE: 06-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 09010/024002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 858/677-1456  
TELEFAX: 858/677-1465  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 956 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-10-121-032-63

Query Match 3.3%; Score 97; DB 9; Length 956;  
Best Local Similarity 20.4%; Pred. No. 3.1;  
Matches 137; Conservative 64; Mismatches 225; Indels 246; Gaps 34;

QY 5 SKDAPNM--DGTSGAGQLVPEANTAEPIS-----MEPVAGAATAAATAGQV 49  
DB 191 SKD-PDNLIIIVGTSNYSQQVDVA-SADPISDTNVAYTLHFAAFNPHDNLNVAQTALDN 248  
QY 50 NM---IDPWIMNYYVQAPQGETISPNWTPGDILFDLQLGPHLNPFLSHLAQMYNGWV-- 104  
DB 249 NVALFVTEW--GTILNTGOGEPDKESTNTWMAFLKEKGIS--HANWSLSDKAFETGTSVYQ 305  
QY 105 GNMKVVLLAGNAFTAGKIIISCI-----PGFAAQNISIAQAATMEPHVIADVRV 154  
DB 306 AGGVSGLSLNKLTASGEIVKNIQWDTETGTQCTTQCTIECIRAMETAQAGDEII 365  
QY 155 LEPIEVLPEEDVRNVLPHNNNDNAPTRMLVCLMYTLPRASGSSGTDPFVIAGRVLTCPSPD 214  
DB 366 IAPGNYNFQDKIQGAFNRS-----VY--LYGSANGNSTNPILILGESATNP-PV 411  
QY 215 FSFL-----FLVP-----PNVEOKT-----KPFVSNPLPLNTL----- 242  
DB 412 FSLGLDYNNGYLLSIEGDYWNIRKIDFKTSGKIVLDNSNGSKLKLNVVHDIGEEAHLRD 471  
QY 243 --SNSRV-----PSLIKSMWVRDGHQVMQVFONG-----RVTLDS 275  
DB 472 GSSNNSIDCCTYNTGRTKPGFEGEGLYVSGDKGQHDYTERACNNNTIENCTVGPNNVTAG 531  
QY 276 --OLOQTTPTSASQLCKIRGSVFHANGNGYNLTE-----LDGSPYH 315

DB 532 VDVKECTMTI-----IRNCVFSABGISENSSDAFIDLKCAYGfVYRNTFNVDGSEV- 584  
QY 316 AFESPAPIGFDPDGLCECDWIMEASPTTQFNTGDVIKOINVKQESAFAPHLGTIOADGLSDV 375  
DB 585 ---INTGVDFLDRG-----TGFNTG---FRNAIFENTY--NLGS----- 615  
QY 376 SYNTNMIAKLGVSPVSDGHRGDVDPWVIPRYGXSTLUEAAQLAPPIYPPCFGEAIVFFMS 435  
DB 616 -----RASEISTARKKOGSP-----EOTHVMDNIRNP-----SV 645  
QY 436 DEPIAHTNCL-----SVPCTIPQEFVTHFVNEQAPTGEAALLHYLDP----- 479  
DB 646 DEPIDSGTENLVNKFPCPDWNIIEPCNPVDE-----TNOAPT-----ISFLSPVNNITLV 693  
QY 480 -----DTHRNLGEEKLYPEGFMTCVPNSS-----GTGPQTLPINGVFFVFSW 521  
DB 694 EGYNLOVEVNAITDADGTIDNVKLYIDNNLVQRINSTYSYKMGHSDSPNTDELNCL----- 747  
QY 522 VSRFYQLKPVGT 533  
DB 748 TECTYTLKAIAAT 759  
RESULT 14  
US-09-738-626-6888  
Sequence 6888, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIJO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6888  
LENGTH: 1344  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6888

Query Match 3.3%; Score 96; DB 9; Length 1344;  
Best Local Similarity 18.4%; Pred. No. 6.3;  
Matches 90; Conservative 57; Mismatches 167; Indels 176; Gaps 22;

QY 10 TNNMGTSAGQLVPEANTAEPISMEPVAGAATAAATAGQVNMIDPWIMNYYVQAPQGET 69  
DB 147 TNNMGSDGFEYVEITNTTAEPIDFSDT-----LNLY-PODEFT 185  
QY 70 ISPNNT-----PGDIL-----FDLQLGPHLNPFLSHLAQMYNGWGNMKVKVLL 113  
DB 186 ----NTNEAVAAEPGDVIIQPKSLVFWIKNCPNDEATAADNAEY----- 228  
QY 114 AGNAFTAGKIIISCTPPGFA---AQNISTAQAATMFPHVI-----ADVRVLEPIEV 160  
DB 229 -GTNLEAGKDLVEISSGGMANGTARCMQIQNTNT--GHIVNRGPNMAGASDVKANEGCLHF 285

```

QY 161 PLED---VRNVLEHNNDNAPTHRLVCMLYTPLRASGSSGTDPFVIAGRVLTJC-PSPDFS 216
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 286 AVDESILLKOTLVGSGAPTFGVYTSQIPNPLSAVIDSSV-PLTIDNTATSIINPAEFT 344
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 217 FLFVPPNVBQKTKPFSPV------LPLNTLSNSRVPSLIKSM 255
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 345 FAFNITDDQVVRATLHVTSAGEAATTINLTEDGCFNNWALPAADLTGK---SWFEVTV 401
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 256 VSRDHGMQVQFONGRVTLQDQ-----LQGT------ 281
    : : : : : : : : : : : : : : : : : : : : : :
Db 402 TATDGFNSVTTEPVRTVDGANTDPLRLNLEENQWVSGTTDVGASDVGDKLELLIDDA 461
    : : : : : : : : : : : : : : : : : : : : : :
QY 282 -----PYSASOLCKIRGSVFHANG -GNGYNLTDELQSPY---HAFESPAPICF 325
    || : : : : || : : || : : || : : || : : || : : || : : || : : || : :
Db 462 PAVTNSLSAAPTFAMEVTQT--DVFFRNGILAGGELRIFDQGYVANTETISTVPVL-- 517
    : : : : : : : : : : : : : : : : : : : : : :
QY 326 PDLGECQWHM-----EASPTQFTGTGVIRKQINVKQESAFAPHLGTTQADG 371
    : : : : : : : : : : : : : : : : : : : : : :
Db 518 -----YHINEDGTLVSVYAGTKAAPEIDLNN--DDQIRNLRLLPDGRILTTPAG 568
    : : : : : : : : : : : : : : : : : : : : : :
QY 372 LSDVSVAINTM 381
    : : : : : : : : : : : : : : : : : : : : : :
Db 569 ISDSNAWLNM 578
    : : : : : : : : : : : : : : : : : : : : : :

```

RESULT- 15

US-09-712-363-158  
; Sequence 158, Application US/09712363  
; Patent No. US20020164588A1

GENERAL INFORMATION:

; GENERAL INFORMATION.  
 ; APPLICANT: Eisenberg, David  
 ; APPLICANT: Rotstein, Sergio H.  
 ; APPLICANT: Marcotte, Edward M.

APPLICANT: MARCOLLE, EDWARD M.  
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND INTERACTIONS OF PROTEINS BY

; TITLE OF INVENTION: INTERAC  
FILE REFERENCE: 07419-032001

FILE REFERENCE: 07419-032001  
CURRENT APPLICATION NUMBER: US/09/712.363

; CURRENT APPLICATION NUMBER: 2000-11-11  
 ; CURRENT FILING DATE: 2000-11-11

: CURRENT FILING DATE: 2000-11-13  
 : PRIOR APPLICATION NUMBER: PCT/US00/022246

;; PRIOR APPLICATION NUMBER: PCI-  
PRIOR FILING DATE: 2000-01-28

; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 604179 531

; PRIOR APPLICATION NUMBER: 607-  
 PRIOR FILING DATE: 2000-03-01

; PRIOR FILING DATE: 2000-02-01  
 ; PRIOR APPLICATION NUMBER: 60/117 844

; PRIOR APPLICATION NUMBER: 60/  
PRIOR FILING DATE: 1989-01-29

; PRIOR FILING DATE: 1999-01-29  
 PRIOR APPLICATION NUMBER: 60/118 206

; PRIOR APPLICATION NUMBER: 60/  
 PRIOR FILING DATE: 1999-03-01

; PRIOR FILING DATE: 1999-02-01  
 PCT APPLICATION NUMBER: 50/136 593

;; PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 1999-03-26

; PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/

; PRIOR APPLICATION NUMBER: 10081319

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/

; PRIOR FILING DATE: 1999-11

NUMBER OF SEQ ID NOS: 292

```
; SOFTWARE: FastSEQ for Windows
```

; SEQ ID NO 158

; LENGTH: 678

TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis

Query Match	3.38:	Score 95.5:	DB 9:	Length 678:
-------------	-------	-------------	-------	-------------

Query Match	3.38;	SCORE 33.3;	DB 3;	length 0.00;
best recall similarity	21 68;	Pred No. 2.5;		

[illegible]

7 DAPTNM-----DGTSGAGQLVPEANTAEPI SMEPVAGAAATAATAGQVNMIDPWIMNN 59

[illegible]

60 YVQAPQCEET---TSPNNTPGDITLFDLQGPHLNPFLSHLAQMY-----NGWVG 105

D <sub>b</sub>	75	FYSNPGSFSTCFARAVKNNLFG-----DLOGG-----STTQQVYKNALVGSAGQHSG	124
Q <sub>y</sub>	106	NW-NVKVLLAGNAFTAG-----KIISICPPGFAAQNISIAQTMPHFVIADVRLIEP	157
D <sub>b</sub>	125	LMRAKELVATATMSGBSKDDVLQAYLNIIYFGRGAYGISAAKSAYFDKPVEQLTVAE-	183
Q <sub>y</sub>	158	IEVPLEDRVNLFHNDNAFTMKLVCMLYTPLRASGSSSGTD-----FVIAGRVLIT	209
D <sub>b</sub>	184	-----GALLAALI-----RRPSTLDPAVDPEGAHARNWNWDGMGVET	220
Q <sub>y</sub>	210	---CPSPDFSELF--LYPPNV--EOKTKP-----FSVPNLPLNTLSNS	245
D <sub>b</sub>	221	KALSPNDRAAQVFDETVPDPDLARENGTKPNGLIERQVTRELELLEFNIDEQTLNI---	276
Q <sub>y</sub>	246	RVPSLIKSMYVRSDHGQMVQFONGRVPTLDGQ-----LQTTPTSASOLCKI---	291
D <sub>b</sub>	277	-----OQLVLT-----TTIDPQARAEKAKVAKYLDGGDPDMRAAVVSI DP	317
Q <sub>y</sub>	292	-RGSVYEHANG--NGYNMLTEL--DGSPHYATESPAP----IGFPDLCEGDWHMEASTP	340
D <sub>b</sub>	318	HNCARAYVYGGDNANGFFDAQAGLQTSSKFVALVAALEQGQ----LG----TQVDS S PL	371
Q <sub>y</sub>	341	TQFTGVDVIKQINWKQESAFAPHLGTIQADCLSDVS VNTN-----MIAKLGWSPVSD-GH	395
D <sub>b</sub>	372	TV-----DGIKITNVBEGEC-----GTCNIAEALKKSLNTSYRMLMLKNGQPQAVADA AH	422
Q <sub>y</sub>	396	RGDVPDWIPRYGSTLTBAEALAPPYPGGEAIVF-----FMSDRPIAHGT-----NGLS	475
D <sub>b</sub>	423	OAGIAS-SFPGPVAHTLSEDKGGPP-----NNGIVLCQYQTRVIDMASAYATA LAASGI Y	475
Q <sub>y</sub>	448	VPCITIQEVTHFVNQEAPTGERAALLHYLDPDTHRNLFGEKL	490
D <sub>b</sub>	476	HP-----PHEVGKVVSANGQV-----LFDASTDANTNGDQRI	506

search completed: January 16, 2003, 10:00:14

Search completed:  
Job time : 22 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:54:55 ; Search time 147 Seconds  
(without alignments)  
2390.337 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNMDTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main.\*  
1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*  
33: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*  
34: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*  
35: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*  
36: /cgn2\_6/ptodata/1/paa/US111\_COMB.pep.\*  
37: /cgn2\_6/ptodata/1/paa/US112\_COMB.pep.\*  
38: /cgn2\_6/ptodata/1/paa/US113\_COMB.pep.\*  
39: /cgn2\_6/ptodata/1/paa/US114\_COMB.pep.\*  
40: /cgn2\_6/ptodata/1/paa/US115\_COMB.pep.\*  
41: /cgn2\_6/ptodata/1/paa/US116\_COMB.pep.\*  
42: /cgn2\_6/ptodata/1/paa/US117\_COMB.pep.\*  
43: /cgn2\_6/ptodata/1/paa/US118\_COMB.pep.\*  
44: /cgn2\_6/ptodata/1/paa/US119\_COMB.pep.\*  
45: /cgn2\_6/ptodata/1/paa/US120\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2896	100.0	545	23	US-09-926-799-1
2	2605.5	90.0	544	6	US-08-273-257-8
3	1979	68.3	544	23	US-09-926-799-4
4	1977.5	68.3	546	23	US-09-926-799-3
5	1950	67.3	530	3	US-07-696-454-3
6	1950	67.3	530	3	US-07-941-365C-3

7	1950	67.3	530	3	US-07-941-365E-3	Sequence 3, Appli
8	1950	67.3	530	7	US-08-386-365-3	Sequence 3, Appli
9	1950	67.3	530	8	US-08-486-049-3	Sequence 3, Appli
10	1938	66.9	530	23	US-09-926-799-2	Sequence 2, Appli
11	1225	42.3	540	23	US-09-926-799-7	Sequence 7, Appli
12	1184.5	40.9	548	6	US-08-273-257-5	Sequence 5, Appli
13	1178.5	40.7	548	23	US-09-926-799-6	Sequence 6, Appli
14	1176.5	40.6	535	6	US-08-273-257-11	Sequence 11, Appl
15	1174.5	40.6	535	23	US-09-926-799-8	Sequence 8, Appli
16	1172.5	40.5	542	23	US-09-926-799-9	Sequence 9, Appli
17	1157.5	40.0	539	23	US-09-926-799-5	Sequence 5, Appli
18	1154.5	39.9	550	23	US-09-926-799-10	Sequence 10, Appl
19	1013	35.0	541	23	US-09-926-799-11	Sequence 11, Appl
20	316	10.9	576	21	US-09-791-537-68473	Sequence 68473, A
21	315	10.9	576	21	US-09-791-537-92267	Sequence 92267, A
22	311	10.7	576	21	US-09-791-537-25489	Sequence 25489, A
23	289	10.0	861	21	US-09-791-537-61752	Sequence 61752, A
24	287.5	9.9	691	21	US-09-791-537-143497	Sequence 143497, A
25	285	9.8	626	1	PCT-US00-15750-7	Sequence 7, Appli
26	285	9.8	770	21	US-09-791-537-146166	Sequence 146166, A
27	282.5	9.8	863	21	US-09-791-537-143865	Sequence 143865, A
28	277	9.6	565	21	US-09-791-537-110580	Sequence 110580, A
29	276	9.5	547	19	US-09-521-738-4	Sequence 4, Appli
30	276	9.5	671	19	US-09-521-738-2	Sequence 2, Appli
31	272.5	9.4	669	20	US-09-617-594-2	Sequence 2, Appli
32	272.5	9.4	669	20	US-09-617-594A-2	Sequence 2, Appli
33	272	9.4	623	1	PCT-US00-15750-4	Sequence 4, Appli
34	272	9.4	668	20	US-09-617-594-4	Sequence 4, Appli
35	272	9.4	668	20	US-09-617-594A-4	Sequence 4, Appli
36	269	9.3	623	1	PCT-US00-15750-2	Sequence 2, Appli
37	262.5	9.1	622	1	PCT-US00-15750-6	Sequence 6, Appli
38	149.5	5.2	878	21	US-09-791-537-151872	Sequence 151872, A
39	147	5.1	878	21	US-09-791-537-151869	Sequence 151869, A
40	143.5	5.0	878	21	US-09-791-537-13397	Sequence 13397, A
41	142.5	4.9	878	21	US-09-791-537-13415	Sequence 13415, A
42	137.5	4.7	912	21	US-09-791-537-19633	Sequence 19633, A
43	135.5	4.7	879	21	US-09-791-537-76207	Sequence 76207, A
44	128.5	4.4	613	21	US-09-791-537-109700	Sequence 109700, A
45	128	4.4	889	21	US-09-791-537-42411	Sequence 42411, A

#### ALIGNMENTS

RESULT 1  
US-09-926-799-1  
; Sequence 1, Application US/09926799  
; GENERAL INFORMATION:  
; APPLICANT: TAKEDA, NAOKAZU  
; APPLICANT: NATORI, KATSURO  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: KAWATA, KUNIO  
; APPLICANT: SATO, TOSHINORI  
; APPLICANT: SATO, SEIYA  
; TITLE OF INVENTION: Detection kit for SRSV  
; FILE REFERENCE: 217039S0XPCT  
; CURRENT APPLICATION NUMBER: US/09/926,799  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: JP 11175928  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: JP 11-175928  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 545  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-09-926-799-1

Query Match 100.0% Score 2896; DB 23; Length 545;

		REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: (415) 543-9600	
		TELEFAX: (415) 543-5043	
		INFORMATION FOR SEQ ID NO: 8:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 544 amino acids	
		TYPE: amino acid	
		TOPOLOGY: linear	
		MOLECULE TYPE: protein	
		US-08-273-257-8	
		Query Match 90.0%; Score 2605.5; DB 6; Length 544;	
		Best Local Similarity 89.2%; Pred. No. 1e-241;	
		Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;	
Qy	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMNNY	60
Db	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMNNY	60
Qy	61	VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA	120
Db	61	VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA	120
Qy	121	GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVLDPVLEDRVNLPHNNDAPTMR	180
Db	121	GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVLDPVLEDRVNLPHNNDAPTMR	180
Qy	181	LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240
Db	181	LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240
Qy	241	TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG	300
Db	241	TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG	300
Qy	301	GNYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGQVTKQINVKQESAF	360
Db	301	GNYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGQVTKQINVKQESAF	360
Qy	361	APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP	420
Db	361	APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP	420
Qy	421	IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEFVTHFVNEQAPTRGEAALLHYLDPD	480
Db	421	IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEFVTHFVNEQAPTRGEAALLHYLDPD	480
Qy	481	THRNLEEFKLYPEGFTMTCVPNSSGTPQTLPLNGVVFVSWSRFYQLKPVGTAGPACRL	540
Db	481	THRNLEEFKLYPEGFTMTCVPNSSGTPQTLPLNGVVFVSWSRFYQLKPVGTAGPACRL	540
Qy	541	GIRRS 545	
Db	541	GIRRS 545	
RESULT 3			
US-09-926-799-4			
; Sequence 4, Application US/09926799			
; GENERAL INFORMATION:			
; APPLICANT: TAKEDA, NAKAZU			
; APPLICANT: NATORI, KATSURO			
; APPLICANT: MIYAMURA, TATSUO			
; APPLICANT: KAMATA, KUNIO			
; APPLICANT: SATO, TOSHINORI			
; APPLICANT: SATO, SEIYA			
; TITLE OF INVENTION: Detection kit for SRSV			
; FILE REFERENCE: 217039USOXPT			
; CURRENT APPLICATION NUMBER: US/09/926,799			
; CURRENT FILING DATE: 2002-03-29			
; PRIOR APPLICATION NUMBER: JP 11175928			
; PRIOR FILING DATE: 1999-06-22			

; PRIOR APPLICATION NUMBER: JP 11-175928  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-09-926-799-4

Query Match 68.3%; Score 1979; DB 23; Length 544;  
Best Local Similarity 66.8%; Pred. No. 3.3e-181;  
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;  
  
Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPIMNNY 60  
Db 1 MMASKDAPTSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPWLNMF 60  
  
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSLAQMYNGVGMKVKVLLAGNAFTA 120  
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSLAQMYNGVGMKVKVLLAGNAFTA 120  
  
Qy 121 GKILISCIIPPGFAAQNISIAQATMPHVIADRVVLEPIEVLDPEDVRNVLFNHNDNAPTMR 180  
Db 121 GXVIVCCVPPGQSTLSAQATLFPVHVIADRVTLDPVEVLDPEDVRNVLHNNDQPTMR 180  
  
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPESVPLNPLN 240  
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPESVPLNPLN 240  
  
Qy 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGQLQGTTPPTSASQLCKIRGSVFHANG 300  
Db 241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCITDQGLGTTTPVSVSQLCKFRGRI--TSG 298  
  
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDDHMEAS--PTQNTGDVIVKQINVKOES 358  
Db 299 QRVNLTLTLDGSPYHAFESPAPIGPFDLGECDDHMEAS--PTQNTGDVIVKQINVKOES 358  
  
Qy 359 AFAPHLGTIQADGLSDVSVNTNIAKLGWSPVSDGHRGCDVDPWIPRYGSTLTEAAQLA 418  
Db 359 QVPHLSSTILD--ENVSSGGYIGTIQWTSPPSDSGGANTFWKIPDYVSSLAESAQLA 416  
  
Qy 419 PPIYPGFGAEALVFFMSDFPIAHGTNGLS----VPCITPQEEVTHFVNEQAPTRGEAALL 474  
Db 417 PAVYPPGGENVIVFMASSIP--GPNQSGSNLVPCLLPQEVITHFISEQAPIQGEAALL 473  
  
Qy 475 HYLDPDTHRNLFGEFKLYPEGFTVCVPNSSGTCGPQTLPLNGVVFVSWVSRRFYQLKPVGTA 534  
Db 474 HYVDPDTNRLNLFGEFKLYPGGYLTCVPNSSSTGPGQQLPLDGVFVFASWSRRFYQLKPVGTA 533  
  
Qy 535 GPA-CRLGIRR 544  
Db 534 GPARGLGVR 544

RESULT 4  
US-09-926-799-3  
; Sequence 3, Application US/09926799  
; GENERAL INFORMATION:  
; APPLICANT: TAKEDA, NAOKAZU  
; APPLICANT: NATORI, KATSURO  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: KAWATA, KUNIO  
; APPLICANT: SATO, TOSHINORI  
; APPLICANT: SATO, SEIYA  
; TITLE OF INVENTION: Detection kit for SRSV  
; FILE REFERENCE: 217039US0XPT  
; CURRENT APPLICATION NUMBER: US/09/926,799  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: JP 11175928  
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: JP 11-175928  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 546  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-09-926-799-3

Query Match 68.3%; Score 1977.5; DB 23; Length 546;  
Best Local Similarity 66.6%; Pred. No. 4.6e-181;  
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;  
  
Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPIMNNY 60  
Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLMEPVAGPTTAVATAGOVNMIDPIMNNF 60  
  
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSLAQMYNGVGMKVKVLLAGNAFTA 120  
Db 61 VQSPQGEFTISPNTPGDILFDLQGLPHLNPFLSLAQMYNGVGMKVMRVRILLAGNAFSA 120  
  
Qy 121 GKILISCIIPPGFAAQNISIAQATMPHVIADRVVLEPIEVLDPEDVRNVLFNHNDNAPTMR 180  
Db 121 GXIIVCCVPPGFTSSSLTIAQATLFPVHVIADRVTLDPVEVLDPEDVRNVLHNNDQPTMR 180  
  
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPESVPLNPLN 240  
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPESVPLNPLN 240  
  
Qy 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGQLQGTTPPTSASQLCKIRGSVFHANG 300  
Db 241 TLSNSRFPSLIOGMLSPDASQVQFQNGRCILIDQLGLTTPATSCQLFRVRGKI--NOC 298  
  
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDDHMEASPT--TQNTGDVIVKQINVKOE-S 358  
Db 299 ARTLNLTEVDGKPFMAFDPAPVGFDFGKCDWHMRISKTNPNTSSGDPMSRVSVQTNVO 358  
  
Qy 359 AFAPHLGTIQADGLSDVSVNTNIAKLGWSPVSDGHRGCDVDPWIPRYGSTLTEAAQLA 418  
Db 359 GVPHLGSITQFDEVENHPTG--DYIGTIEWISQPSPTPGTDINLWEIPDYVSSLSQAANLA 417  
  
Qy 419 PPIYPGFGAEALVFFMSDFPIAHGTNGLS----VPCITPQEEVTHFVNEQAPTRGEAALL 474  
Db 418 PVVFPFGFEALVYFVSAP--GPNRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALL 474  
  
Qy 475 HYLDPDTHRNLFGEFKLYPEGFTVCVPNSSGTCGPQTLPLNGVVFVSWVSRRFYQLKPVGTA 534  
Db 475 HYVDPDTNRLNLFGEFKLYPGGYLTCVPNGVAGPQQLPLNGVFLFVSWVSRRFYQLKPVGTA 534  
  
Qy 535 GPA-CRLGIRR 544  
Db 535 STARSRLGVR 545

RESULT 5  
US-07-696-454-3  
; Sequence 3, Application US/07696454  
; GENERAL INFORMATION:  
; APPLICANT: Estes, Mary K.  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y.  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses.  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; STREET: Patent Department, Fulbright & Jaworski  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA

```

; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect converted to DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/696,454
; FILING DATE: 19910506
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5179CIP-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-3634
; TELEFAX: (713) 651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-696-454-3

Query Match 67.3%; Score 1950; DB 3; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMWASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMWASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
QY 121 GKIIISCIIPPGFAAQNISTAQATMPEPHVIADRVRLPEIEVPLEDVNRVLFHNND-NAPTM 179
DB 121 GKIIISCIIPPGFAAQNISTAQATMPEPHVIADRVRLPEIEVPLEDVNRVLFHNND-NAPTM 179
QY 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFESVPLNPL 239
DB 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFESVPLNPL 239
QY 181 RLVCMLYTLRLTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
DB 181 RLVCMLYTLRLTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
QY 240 NTLNSRVPSLKMVSRRDHQGMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVFHAN 299
DB 240 NTLNSRVPSLKMVSRRDHQGMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVFHAN 299
QY 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT---S 294
DB 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT---S 294
QY 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTQFNTGDVTKQINVKQESA 359
DB 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTQFNTGDVTKQINVKQESA 359
QY 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQFGHSSQTYDVTPT 351
DB 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQFGHSSQTYDVTPT 351
QY 360 FAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
DB 360 FAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
QY 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNYGSSITEATHLAP 407
DB 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNYGSSITEATHLAP 407
QY 420 FYPPPGFEALVFMSPDPIAHNGTNGLSVPTCIQEFVTHVNEQAQTRCEAAALLHYLDP 479
DB 420 FYPPPGFEALVFMSPDPIAHNGTNGLSVPTCIQEFVTHVNEQAQTRCEAAALLHYLDP 479
QY 408 SVYPPGFEVLVFFMSKMP---GFGAYNLCLLPQEYISHLASQAPVTGGAALLHYVDP 464
DB 408 SVYPPGFEVLVFFMSKMP---GFGAYNLCLLPQEYISHLASQAPVTGGAALLHYVDP 464
QY 480 DTHRLNGEFLKYPGFMTCVPSNDSGTGQPTLPINGVFVFSWVSRYQLKPVGTAGPA-C 538
DB 480 DTHRLNGEFLKYPGFMTCVPSNDSGTGQPTLPINGVFVFSWVSRYQLKPVGTAGPA-C 538
QY 465 DTHRLNGEFLKYPGFMTCVPSNDSGTGQPTLPINGVFVFSWVSRYQLKPVGTASSARG 524
DB 465 DTHRLNGEFLKYPGFMTCVPSNDSGTGQPTLPINGVFVFSWVSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
DB 525 RLGLRR 530

RESULT 6
US-07-941-365C-3
```

```

; Sequence 3. Application US/07941365C
; GENERAL INFORMATION:
; APPLICANT: Matson, David O
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski Patent Dept
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/941,365C
; FILING DATE: 19920908
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-651-3634
; TELEFAX: 713-651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-941-365C-3
```

```

Query Match 67.3%; Score 1950; DB 3; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMWASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMWASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
QY 121 GKIIISCIIPPGFAAQNISTAQATMPEPHVIADRVRLPEIEVPLEDVNRVLFHNND-NAPTM 179
DB 121 GKIIISCIIPPGFAAQNISTAQATMPEPHVIADRVRLPEIEVPLEDVNRVLFHNND-NAPTM 179
QY 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFESVPLNPL 239
DB 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFESVPLNPL 239
QY 181 RLVCMLYTLRLTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
DB 181 RLVCMLYTLRLTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
QY 240 NTLNSRVPSLKMVSRRDHQGMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVFHAN 299
DB 240 NTLNSRVPSLKMVSRRDHQGMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVFHAN 299
QY 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT---S 294
DB 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGT---S 294
QY 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTQFNTGDVTKQINVKQESA 359
DB 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTQFNTGDVTKQINVKQESA 359
QY 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQFGHSSQTYDVTPT 351
DB 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQFGHSSQTYDVTPT 351
QY 360 FAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
DB 360 FAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
QY 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNYGSSITEATHLAP 407
DB 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNYGSSITEATHLAP 407
```



Thu Jan 16 15:59:22 2003

```
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;
QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNNY 60
Db 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
QY 180 RLVCMLYTLPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPL 239
Db 180 RLVCMLYTLPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPL 239
QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
Db 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
QY 295 NGTVINLTDELGTPEHPFEGPAPIGFPLDGLGCDWHIN---MTQFGHSSQTYDVTTPDT 351
Db 295 NGTVINLTDELGTPEHPFEGPAPIGFPLDGLGCDWHIN---MTQFGHSSQTYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTTEAAQLAP 419
Db 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTTEAAQLAP 419
QY 352 FVPHLGSIQANGIG---SGNVGVLSWISPSHPSGSGQVLDLWKIPNYGSSITEATHLAP 407
Db 352 FVPHLGSIQANGIG---SGNVGVLSWISPSHPSGSGQVLDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGAIVFFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 420 PIYPGFGAIVFFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
QY 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDP 464
Db 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDP 464
QY 480 DTHRNLFGEFKLYPEGFMTCPVNSSGTGPQTLPIGVFVFSWVSRYQLKPVGTAGPA-C 538
Db 480 DTHRNLFGEFKLYPEGFMTCPVNSSGTGPQTLPIGVFVFSWVSRYQLKPVGTAGPA-C 538
QY 465 DTGRNLGEFKAYPDGFLTCVPNGASSGQQLPIGVFVFSWVSRYQLKPVGTASSARG 524
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGQQLPIGVFVFSWVSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
Db 525 RLGLRR 530
RESULT 10
US-09-926-799-2
; Sequence 2, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039USXPCT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
```

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-049-3
Query Match 67.3%; Score 1950; DB 8; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;
QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNNY 60
Db 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVPLEDVRNVLHNNND-NAPTM 179
QY 180 RLVCMLYTLPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPL 239
Db 180 RLVCMLYTLPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPL 239
QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
Db 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPTLNLPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
QY 295 NGTVINLTDELGTPEHPFEGPAPIGFPLDGLGCDWHIN---MTQFGHSSQTYDVTTPDT 351
Db 295 NGTVINLTDELGTPEHPFEGPAPIGFPLDGLGCDWHIN---MTQFGHSSQTYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTTEAAQLAP 419
Db 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTTEAAQLAP 419
QY 352 FVPHLGSIQANGIG---SGNVGVLSWISPSHPSGSGQVLDLWKIPNYGSSITEATHLAP 407
Db 352 FVPHLGSIQANGIG---SGNVGVLSWISPSHPSGSGQVLDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGAIVFFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 420 PIYPGFGAIVFFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
QY 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDP 464
Db 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDP 464
QY 480 DTHRNLFGEFKLYPEGFMTCPVNSSGTGPQTLPIGVFVFSWVSRYQLKPVGTAGPA-C 538
Db 480 DTHRNLFGEFKLYPEGFMTCPVNSSGTGPQTLPIGVFVFSWVSRYQLKPVGTAGPA-C 538
QY 465 DTGRNLGEFKAYPDGFLTCVPNGASSGQQLPIGVFVFSWVSRYQLKPVGTASSARG 524
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGQQLPIGVFVFSWVSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
Db 525 RLGLRR 530
RESULT 10
US-09-926-799-2
; Sequence 2, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039USXPCT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
```

;> PRIOR APPLICATION NUMBER: JP 11-175928  
;> PRIOR FILING DATE: 1999-06-22  
;> NUMBER OF SEQ ID NOS: 34  
;> SOFTWARE: PatentIn version 3.1  
;> SEQ ID NO 2  
;> LENGTH: 530  
;> TYPE: PRT  
;> ORGANISM: ARTIFICIAL SEQUENCE  
;> FEATURE:  
;> OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-09-926-799-2

Query Match 66.9%; Score 1938; DB 23; Length 530;  
Best Local Similarity 67.0%; Pred. No. 2.9e-177;  
Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMEPVAGATAATAAGQVNMIDPWIMNNY 60  
DB 1 MMASKADTSSVDGASGACOLVPEVNASDPLANDPVAGSTAVATAGQVNPIDPWIIINF 60  
QY 61 VQAPQGEFTISPNNTPGDILFDLQLGPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120  
DB 61 VQAPQGEFTISPNNTPGGVLFDSLGLPHLNPFLSHLAQMYNGVGMKVKVIMLAGNAFTA 120  
QY 121 GKIIISCIPPGFAAQNISIAQATMPHVIADVRVLEPIEVPLEDVRNLFHND-NAPTM 179  
DB 121 GKIIISCIPPGFSGHNLITIAQATLPHVIAVRLTDPDIEVPLEDVRNLFHNDNRQOTM 180  
QY 180 RLVCMLYTLPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPFSVNLPL 239  
DB 181 RLVCMLYTLPLRTGGGTG--DSFVAGRVMTCPSPDFNFLLFVPPVVEQKTRPTLPNLPL 238  
QY 240 NTLNSRVPSLIKSMVSRDHQOMVQFNGRVTLDGLOLQGTTPTSASOLCKIRGSVFHAN 299  
DB 239 SLSLSRAPLPSIGMGISPDNVQSVQFNGRCTLDGRLVGTTPVSLSHVAKIRGT---S 294  
QY 300 GNGYNTLDELSPYHAFESPAPIGPPDILGECDMHMEASPTQFNTGNDVIKQINVKQESA 359  
DB 295 NQTVINLTLDGTTFHFFEGPAPIGPPDILGCGDWHIN---MTQFGHSSQTYQVDVDTPTD 351  
QY 360 FAPHILGTIQADGLSDVSNTNIAKLGWSPVSDGHRGVDWPKVPRYSGTSLTEAAQLAP 419  
DB 352 FPHLGSIQANGIG----SGNTIGVLSWSPSPSHSGSGSDLVKWKIPNYGSSITEATHLAP 407  
QY 420 PIYPGFGGEAIVFEMSDFFIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
DB 408 SVYPGFGGEVLVFNESKIP--GPGAYSLPCLLPQEIYISHLASEQAAPTVEAALLHYVDP 464  
QY 480 DTHRNILGEKLYPEGMTCVPNSSGTPQTLPIGVFVSVWSRYQLKPVGTAGPA-C 538  
DB 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPOQLPIGVFVSVWSRYQLKPVGTASSARG 524  
QY 539 RLGIIR 544  
DB 525 RLGLRR 530

RESULT 11  
US-09-926-799-7  
;> Sequence 7, Application US/09926799  
;> GENERAL INFORMATION:  
;> APPLICANT: TAKEDA, NAKAZU  
;> APPLICANT: NATORI, KATSURO  
;> APPLICANT: MIYAMURA, TATSUO  
;> APPLICANT: KAMATA, KUNIO  
;> APPLICANT: SATO, TOSHINORI  
;> APPLICANT: SATO, SEIYA  
;> TITLE OF INVENTION: Detection kit for SRV  
;> FILE REFERENCE: 217039USOXPT  
;> CURRENT APPLICATION NUMBER: US/09/926,799  
;> CURRENT FILING DATE: 2002-03-29  
;> PRIOR APPLICATION NUMBER: JP 11175928  
;> PRIOR FILING DATE: 1999-06-22

;> PRIOR APPLICATION NUMBER: JP 11-175928  
;> PRIOR FILING DATE: 1999-06-22  
;> NUMBER OF SEQ ID NOS: 34  
;> SOFTWARE: PatentIn version 3.1  
;> SEQ ID NO 7  
;> LENGTH: 540  
;> TYPE: PRT  
;> ORGANISM: ARTIFICIAL SEQUENCE  
;> FEATURE:  
;> OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-09-926-799-7

Query Match 42.3%; Score 1225; DB 23; Length 540;  
Best Local Similarity 46.7%; Pred. No. 2.2e-108;  
Matches 261; Conservative 89; Mismatches 173; Indels 36; Gaps 14;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMEPVAGATAATAAGQVNMIDPWIMNNY 60  
DB 1 MMASKADATPSNDGAAG---LVPESNN-EAMALEPVVGASLAAPVTGQTNIIDPWIRTNF 56  
QY 61 VQAPQGEFTISPNNTPGDILFDLQLGPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120  
DB 57 VQAPNGEFTVSPNSPCEILVNLGPELNPYLALHARMYNGAGGMEVQVMLAGNAFTA 116  
QY 121 GKIIISCIPPGFAAQNISIAQATMPHVIADVRVLEPIEVPLEDVRNLFH-NNDNAPTM 179  
DB 117 GKIIIFAAVPPYFVENLSPSQITMTMPHVIDVRLTLEPVLPMDPDVRSTLFHFQKDEPKM 176  
QY 180 RLVCMLYTLPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPFSVNLPL 239  
DB 177 RLVCMLYTLPLRSNG--SGDDVFTVSCRILTRPSPEEDTYLVPPVTVESKTKPTLPVLTL 234  
QY 240 NTLNSRVPSLIKSMVSRDHQOMVQFNGRVTLDGLOLQGTTPTSASOLCKIRGSVFHAN 299  
DB 235 GELNSRFPPLSDIEWTSPNESIVQVQNGRVTLDELLGTTQLQACNCSIRGKVTGQV 294  
QY 300 GNGY----NLTELDGSPYHAFES-PAPIGFPDL-GECDMH-----EASPTTQFNTGD 347  
DB 295 PSEQHMNLEITNLNGTQDPTDDVPALGVDPDFAGEVFGVLSQRNRESNPANRAHDV 354  
QY 348 VIKQINVKQESAFAPHLGTIQAD--GLSDVSNTNIAKLGWSPVSDGHRGVDWPKV 405  
DB 355 V-----ATYSKYTPKLGVLQVGTWNTNDVNOPTKFTPIG-LNEVANGHR--FEQWTL 406  
QY 406 RYGSTLTEAAQLAPPIYPGFGGEAIVFEMSDFFIAHGTNGLSVPCITPQEFVTHFVNEQA 465  
DB 407 RYSGALTLMNNLAPAVAPLEPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHFQESA 466  
QY 466 PTRGEAALLHYLDPDTHRNILGEKLYPEGMTCVPNSSGTPQTLPIGVFVSVWSRYQLK 525  
DB 467 PSLGDVALVRYVNPDTGRVLFQAKLHKGGLTV--SSTSTGPPVVVPANGYFKFDSVNOF 524  
QY 526 YQLKPVGTAGPACRLGIIR 544  
DB 525 YSLAPMGTGN-----GRRR 538

RESULT 12  
US-08-273-257-5  
;> Sequence 5, Application US/08273257  
;> GENERAL INFORMATION:  
;> APPLICANT: LEW, Judy F.  
;> APPLICANT: GREEN, Kim Y.  
;> APPLICANT: VALDESUSO, Jose  
;> TITLE OF INVENTION: Calicivirus capsid genes and their uses  
;> NUMBER OF SEQUENCES: 36  
;> CORRESPONDENCE ADDRESS:  
;> ADDRESSEE: Townsend and Townsend Kourie and Crew  
;> STREET: Steuart Street Tower, One Market Plaza  
;> CITY: San Francisco  
;> STATE: California  
;> COUNTRY: US  
;> ZIP: 94105-1493

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,257
; FILING DATE: 11-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-209
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-273-257-5

Query Match 40.9%; Score 1184.5; DB 6; Length 548;
Best Local Similarity 46.4%; Pred. No. 1.8e-104;
Matches 262; Conservative 79; Mismatches 181; Indels 43; Gaps 15;

QY 1 MMASKD-AFTNMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGOVNMDIPMNN 59
Db 1 MKMASNDAAFSN----DGAACLVPEINN-EAMALEPVAGAAIAAPLTGQNIIDPIMNN 55

QY 60 YVQAPQGEFTISPNNTPCDILFDLQGLPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFT 119
Db 56 FVQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFT 115

QY 120 AGKIIISCIPPGFAAQNISIAQTAMFPHVIADRVLEPIEVPLEDVRNVLFHNDNAPT- 178
Db 116 AGKIIFAAIPPNPFDNLNSAAQITMCPHVIVDVRQLEPVLNLPMPDVRNFFHNGSDSR 175

QY 179 MRLVCMLYTLRASGSSGTDPPFIAGRVLTCPSPDFSEFLVPPNVEOKTKPFSVNP 238
Db 176 LRLAAMLYTLPLRA--NNSGDDVFTVSCRVLTRPSDFSEFLVPPTVESKTKPFTLPILT 233

QY 239 NLTLSNRVPSLIKSMVSRDHQGVQFQNGRVTLDCQLQGTTPTSASOLCKIRGSVFHA 298
Db 234 ISEMSNRFPVPIDSLHTSPTEINIVVQCQNGRVTLDCGLMGTTLQLLPSQICAFRGTLTRS 293

QY 299 NGG-----NGY---NLTELDGSPYHAFES-PAPIGFPDL-GECDWHMEASPT 340
Db 294 TSRSADQADTPTPLRFNYWHVQLDNLNGTTPYDPAEDIPALGTPDFRGKV----FGVA 348

QY 341 TQFNTGDVIKQINVKQESA---FAPHLGTIOADGLSDSVSVNTNMIAKLGWSPVSDG--H 395
Db 349 SQRNPDSTTRAHEAKVDTTSGRTFKLGSLEITTESD-DFDQNPQTKP---TPVGVGVDN 404

QY 396 RGVDVPIPRYGSTLTEAAQLAPPIYPGGEAIVFMSDEPIAHGCTNGLSVPCITPOE 455
Db 405 EAEFOQWLSLPNTSGOFTTHMNLAPAVAPNPFGEQLLFFRSQLPSSGSGRNGVLDCLVPQE 464

QY 456 FVTHFVNECAPTRGEAALLHYLDPDTHRNLFGEFKLYPEGFMTCPVNSSGTGQTLPINGV 515
Db 465 WVOHFQESAPAQTOVALVRYVNPDTGRVLFELAKHLKLGEMTIANN--GDSPTVPPNGY 522

QY 516 FVEVSWRSRYOLKPVGTAGPACRL 540
Db 523 FRESWNPFTYTLAPMGTGNGRRRI 547

RESULT 13
US-09-926-799-6
; Sequence 6, Application US/09926799

; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection Kit for SRSV
; FILE REFERENCE: 217039USOXPCT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 548
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
; US-09-926-799-6

Query Match 40.7%; Score 1178.5; DB 23; Length 548;
Best Local Similarity 45.7%; Pred. No. 6.8e-104;
Matches 258; Conservative 82; Mismatches 183; Indels 41; Gaps 14;

QY 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGOVNMDIPWIMNNY 60
Db 1 MKMASNDAAPNDDGAAG---LVPEINN-EAMALDPVAGAAIAAPLTGQNIIDPWINNF 56

QY 61 VQAPQGEFTISPNNTPCDILFDLQGLPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFTA 120
Db 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116

QY 121 GKIIISCIPPGFAAQNISIAQTAMFPHVIADRVLEPIEVPLEDVRNVLFHNDNAPT-M 179
Db 117 GKIIIFAAIPPNPFDNLNSAAQITMCPHVIVDVRQLEPVLNLPMPDVRNFFHNGSDSRL 176

QY 180 RLVCMLYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSEFLVPPNVEOKTKPFSVNP 239
Db 177 RLIAMLYTLPLRA--NNSGDDVFTVSCRVLTRPSDFSEFLVPPTVESKTKPFTLPILT 234

QY 240 NLTLSNRVPSLIKSMVSRDHQGVQFQNGRVTLDCQLQGTTPTSASOLCKIRGSVFHAN 299
Db 235 SEMSNRFPVPIESLHTSPTEINIVVQCQNGRVTLDCGLMGTTLQLLPSQICAFRGVLTTRST 294

QY 300 GG-----NGY---NLTELDGSPYHAFES-PAPIGFPDL-GECDWHMEASPTT 341
Db 295 SRASDAQADTATPLRFNYWHVQLDNLNGTTPYDPAEDIPGCLGTPDFRGKV-----FGVAS 349

QY 342 QFNTGDVIKQINVKQESA---FAPHLGTIOADGLSDSVSVNTNMIAKLGWSPVSDG--HR 396
Db 350 QRNLDSITTRAHEAKVDTTAGRTFKLGSLEISTDSD-DFDQNPQTKF---TPVGIGVDNE 405

QY 397 GDVDPWIPRYGSTLTEAAQLAPPIYPGGEAIVFMSDEPIAHGCTNGLSVPCITPOEF 456
Db 406 AEFQWLSLPNTSGOFTTHMNLAPAVAPNPFGEQLLFFRSQLPSSGSGRNGVLDCLVPQEW 465

QY 457 VTHFVNECAPTRGEAALLHYLDPDTHRNLFGEFKLYPEGFMTCPVNSSGTGQTLPINGVF 516
Db 466 VQHFYQESAPAQTOVALVRYVNPDTGKVLFEAKHLKLGEMTIANN--GDSPTVPPNGYF 523

QY 517 FVEVSWRSRYOLKPVGTAGPACRL 540
Db 524 RFESWNPFTYTLAPMGTGNGRRRI 547

RESULT 14
US-08-273-257-11
; Sequence 11, Application US/08273257
```



; GENERAL INFORMATION:  
 ; APPLICANT: LEW, Judy F.  
 ; APPLICANT: GREEN, Kim Y.  
 ; APPLICANT: VALDESUSO, Jose  
 ; TITLE OF INVENTION: Calicivirus capsid genes and their uses  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew  
 ; STREET: Steuart Street Tower, One Market Plaza  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: US  
 ; ZIP: 94105-1493  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/273,257  
 ; FILING DATE: 11-JUL-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bastian, Kevin L.  
 ; REGISTRATION NUMBER: 34,774  
 ; REFERENCE/DOCKET NUMBER: 15280-209  
 ; REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 543-9600  
 ; TELEFAX: (415) 543-5043  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 535 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-273-257-11

Query Match 40.6%; Score 1176.5; DB 6; Length 535;  
 Best Local Similarity 45.8%; Pred. No. 1e-103;  
 Matches 254; Conservative 88; Mismatches 181; Indels 31; Gaps 12;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNNY 60  
 DB 1 MKMASNDAAAPSNDGAAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNVIDPWIRMF 56  
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120  
 DB 57 VQAPNGEFTVSPRNSPGEILLNLELGPENLPFLSHLSRMYNGYAGGVEQVLLAGNAFTA 116  
 QY 121 GKIIISCIPPGFAAQNISIAQATMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 179  
 DB 117 GKLVFAAIPHPPLENLSFGQITMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 176  
 QY 180 RLVCMLYTLPLRASGSSGSDTDPFVIAGRVLTCPSPDFSFLLVPPNVEQKTKPFSVPLNPL 239  
 DB 177 RLVMAMLYTLPLRSNG--SGDDVFTVSCRVLTRPSDPDFNLVPPVTESKTKPTPLITI 234  
 QY 240 NTLNSRVPFLSLKSMVSRDHGOMQVQNGRVTLDCOLOGTTPTSASOLCKIRGSVFHAN 299  
 DB 235 GELTNSRFPVPIDELYTSNEGVIQVQNGRSTLDGELLGTLQLVPSNICALRGRI-NAQ 293  
 QY 300 GGNYN-----LTELDGSYPHAFES-PAPIGPDLGECOMHMEASPTTQFNGDVIKQ-- 351  
 DB 294 VPDHQQHWNLOVNTNGTFEDTEDYPAELCTDFLANIYGV----TSORNPNNTCRAHD 349  
 QY 352 -INVQKESAFAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGVDPPWIPRYGST 410  
 DB 350 GVLATNSPKFTPKLGSVILGTWESDLNQTRE---TPVGLFNTDFDQWALPSYSGR 406  
 QY 411 LTEAQAQAPPIYPPGGEAIVFMSDFPIAHGTNGLSVPCCTTPOEFTVTHFVNEQAPTRGE 470  
 DB 407 LTLNMNLAPSVSPFLPFGQLLFFRSHIPLKGGTSDGAIDCLLPQEWIQHYESAPAAATD 466

QY 471 AALLHYLDPDTHRLNGEFLKYPGFWMTCPVNSSGTGPOTLPINGVFVSWSRFYQLKP 530  
 DB 467 VALIRTNPDTPGRVLPEAKLHQGFITVA--NSGSRPIVPPNGYFRFDSWVNOFSLAP 524  
 QY 531 VGTAGPACRLGIRR 544  
 DB 525 MGTGN-----GRRR 533  
 RESULT 15  
 US-09-926-799-8  
 ; Sequence 8, Application US/09926799  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TAKEDA, NAOKAZU  
 ; APPLICANT: NATORI, KATSURO  
 ; APPLICANT: MIYAMURA, TATSUO  
 ; APPLICANT: KAMATA, KUNIO  
 ; APPLICANT: SATO, TOSHINORI  
 ; APPLICANT: SATO, SEIYA  
 ; TITLE OF INVENTION: Detection kit for SRSV  
 ; FILE REFERENCE: 217039US0XPCT  
 ; CURRENT APPLICATION NUMBER: US/09/926,799  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: JP 11175928  
 ; PRIOR FILING DATE: 1999-06-22  
 ; PRIOR APPLICATION NUMBER: JP 11-175928  
 ; PRIOR FILING DATE: 1999-06-22  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 8  
 ; LENGTH: 535  
 ; TYPE: PRT  
 ; ORGANISM: ARTIFICIAL SEQUENCE  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE  
 ; US-09-926-799-8

Query Match 40.6%; Score 1174.5; DB 23; Length 535;  
 Best Local Similarity 45.0%; Pred. No. 1.6e-103;  
 Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDPWIMNNY 60  
 DB 1 MKMASNDAAAPSNDGAAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNVIDPWIRNF 56  
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120  
 DB 57 VQAPNGEFTVSPRNSPGEILLNLELGPENLPFLSHLSRMYNGYAGGVEQVLLAGNAFTA 116  
 QY 121 GKIIISCIPPGFAAQNISIAQATMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 179  
 DB 117 GKLVFAAIPHPPLENLSFGQITMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 176  
 QY 180 RLVCMLYTLPLRASGSSGSDTDPFVIAGRVLTCPSPDFSFLLVPPNVEQKTKPFSVPLNPL 239  
 DB 177 RLVMAMLYTLPLRSNG--SGDDVFTVSCRVLTRPSDPDFNLVPPVTESKTKPTPLITI 234  
 QY 240 NTLNSRVPFLSLKSMVSRDHGOMQVQNGRVTLDCOLOGTTPTSASOLCKIRGSVFHAN 299  
 DB 235 GELTNSRFPVPIDELYTSNEGVIQVQNGRSTLDGELLGTLQLVPSNICALRGRI-NAQ 294  
 QY 300 GGNYN-----LTELDGSYPHAFES-PAPIGPDLGECOMHMEASPTTQFNGDVIKQ-- 346  
 DB 295 VPDHQQHWNLOVNTNGTFEDTEDYPAELCTDFLANIYGV----TSORNPNNTCRAHD 348  
 QY 352 -INVQKESAFAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGVDPPWIPRYGST 406  
 DB 349 DAVIATN---SAKFTPKLGAIQIGTWEEEDVHINQTKP---TPVGLFENEGRNQWTLN 402  
 QY 407 YGSLTLEAQAQAPPIYPPGGEAIVFMSDFPIAHGTNGLSVPCCTTPOEFTVTHFVNEQAP 466  
 DB 403 YSGALTLMNLAPSVSPFLPFGQLLFFRSHIPLKGGVADVDICLLPQEWIQHYESAP 462



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:55:20 ; Search time 20 Seconds  
(without alignments)  
1990.084 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDGTSGAGQ.....YOLKPVGTAGPACRLGTRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 271938 seqs, 73030588 residues

Total number of hits satisfying chosen parameters: 271938

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1950	67.3	530	US-10-314-739-3	Sequence 3, Appli
2	272.5	9.4	669	US-10-209-507-2	Sequence 2, Appli
3	272	9.4	668	US-10-209-507-4	Sequence 4, Appli
4	109	3.8	10431	US-60-427-045-310	Sequence 310, App
5	106.5	3.7	931	PCT-US02-33645-26	Sequence 26, Appl
6	103.5	3.6	1210	US-10-258-106-1	Sequence 1, Appli
7	102.5	3.5	1210	US-10-293-017-66	Sequence 66, Appl
8	100	3.5	6879	US-60-419-463-26	Sequence 26, Appl
9	99	3.4	4961	PCT-US02-10366-64	Sequence 64, Appl
10	98	3.4	490	US-10-258-951-67	Sequence 67, Appl
11	97.5	3.4	1328	PCT-US02-39126-6	Sequence 6, Appli
12	97.5	3.4	2126	US-10-052-648A-39	Sequence 39, Appl
13	97	3.3	584	US-09-724-676-91417	Sequence 91417, A
14	97	3.3	584	US-09-724-676A-91417	Sequence 91417, A
15	97	3.3	645	US-09-724-676-91423	Sequence 91423, A
16	97	3.3	645	US-09-724-676A-91423	Sequence 91423, A
17	96.5	3.3	903	US-09-724-676-94702	Sequence 94702, A
18	96.5	3.3	903	US-09-724-676A-94702	Sequence 94702, A
19	96.5	3.3	941	US-09-724-676-94682	Sequence 94682, A
20	96.5	3.3	941	US-09-724-676A-94682	Sequence 94682, A
21	96.5	3.3	943	US-09-724-676-94698	Sequence 94698, A
22	96.5	3.3	943	US-09-724-676-94699	Sequence 94699, A
23	96.5	3.3	943	US-09-724-676-94700	Sequence 94700, A
24	96.5	3.3	943	US-09-724-676-94701	Sequence 94701, A
25	96.5	3.3	943	US-09-724-676A-94698	Sequence 94698, A
26	96.5	3.3	943	US-09-724-676A-94699	Sequence 94699, A

ALIGNMENTS

RESULT 1

US-10-314-739-3

; Sequence 3, Application US/10314739

; GENERAL INFORMATION:

; APPLICANT: Estes, Mary K

; Jiang, Xi

; Graham, David Y

; TITLE OF INVENTION: Methods and Reagents to Detect and

; Characterize Norwalk and Related Viruses

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 801 Pennsylvania Ave., N.W.

; CITY: Washington, D.C.

; STATE: <Unknown>

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/10/314,739

; APPLICATION NUMBER: US/10/314,739

; FILING DATE: 09-Dec-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,049

; FILING DATE: June 7, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Davis, Peter

; REGISTRATION NUMBER: 36,119

; REFERENCE/DOCKET NUMBER: 311.023

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-662-0200

; TELEFAX: 202-662-4643

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 530 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-314-739-3

Query Match 67.3%; Score 1950; DB 6; Length 530;  
Best Local Similarity 67.0%; Pred. No. 1e-160;  
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

Sequence 94700, A  
Sequence 94701, A  
Sequence 94707, A  
Sequence 94707, A  
Sequence 94679, A  
Sequence 94678, A  
Sequence 94679, A  
Sequence 94680, A  
Sequence 94681, A  
Sequence 94678, A  
Sequence 94679, A  
Sequence 94680, A  
Sequence 94681, A  
Sequence 94687, A  
Sequence 94687, A  
Sequence 94703, A  
Sequence 94704, A  
Sequence 94705, A  
Sequence 94706, A  
Sequence 94703, A

Db	126	DGDSSI--TTPEQGTIVGGVIAEPSPAQMAATAADAATGKSVDEW-----SSFSFHF	174
Qy	71	-----SPNNTPCDILFDLOLGHLPNLPFLSHLAQMYNGWGNMKVKVLLAGNAFTAGKLI	125
Db	175	TSVNWSTSEFQGIKILFKQSLUGPLNLPYLEHLSKLYVWMSGVDVREFISGSGVEGKLA	234
Qy	126	SCIPPGF-AAQNT:SIQAQMFPHIVIAADVRLVLEPIEPLDVRNVLFHNNNDNAPTMRKLV	184
Db	235	IVPPGVDVPQVST:MLQ---YPHVLFDAQRQVEPVIFSIPLRSTLYLHMSD:TDTTSLV	291
Qy	185	LY-----TPLRASGSSGTDFFVIAGRVLCPSPDFSLFVLPNPNVEQKTKPSPVNLPL	240
Db	292	YNDLINPYANDSNSG---IVT--VETKPGDFKFLHLKPPG-----S31	331
Qy	241	TLSNRVPS-LI---KSMVYSRHDGQWVQ-----FONGRVTL:DGLOQGTPTTSASOL	288
Db	332	MLTRG:TPSOLIPKSSLSLWIGNRWSDITDFVRPFVQANR-HFEDFN-QETAGWSTPFR	389
Qy	289	CKIRGSVFHANG---GNG-----YNTEL:DGSPY---314	314
Db	390	RPIVITTSSENGSKLGTGVATDIYVPGIPDGWPDITIGEELTPAGDY:SIITNGSGNDIATA	449
Qy	315	HAFESPAPI-----GFPLDGECD-----WHMEASPTQFNT-----GDV:IKQINVKOESAF	360
Db	450	NAYDSADVITNTTFRGMVYICGALQRAWGDKK:STAFITAIKEGNTLKPSTWIMDKL	509
Qy	361	A-----PHLGTQADGLSDVSVNTNMIKLGWV-----SPVSDGHRGDVDWVPIRYG:STLT	412
Db	510	AVYQDTHVG-----RDVQTSDDTLALGYTGGEQAIGSNRDSVVRISMLPETCAR--	560
Qy	413	EAALQAPPIYPGPGFEATVFFMSDFPIAHGTINGLSV:PCTIPQBEVTHFVNEQAPT:GEAA	472
Db	561	-----CGNHPIFYKNSIKLGYVLRSIDV-----FNSQILHTSRQLS	596
Qy	473	LLHYLDP	479
Db	597	LNHYLLP	603

```

RESULT 3
US-10-209-507-4
; Sequence 4, Application US/10209507
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR
; FILE REFERENCE: 4543113-3151-2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PrT
; ORGANISM: Feline calicivirus
US-10-209-507-4

```

Query Match	9.4%;	Score 272;	DB 6;	Length 668;
Best Local Similarity	31.0%;	Pred: 4.5e-15;		
Matches 76; Conservative		66; Mismatches 77;	Indels 56;	Gaps

  

QY	23	PEANT-----AP-ISM	EVAGANTAAATAGVNNIDP	IMNNYVQAQGGFTI----	70
Db	133	PEOTLVGGVIAEPNAQMS	AVADVKGKVDSE-----	W-----EAFSEHTSWN	17
QY	71	SPNPTPCDILFDLQGLPH	NLPFLSHLAQMYNGWGNM	KVKVLLACNATAGKIIIS	12

[illegible]

```

RESULT 2
US-10-209-507-2
; Sequence 2, Application US/10209507
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

```

Query Match 9.4%; Score 272.5; DB 6; Length 669;  
Best Local Similarity 23.4%; Pred. No. 4.1e-15;  
Matches 128; Conservative 76; Mismatches 194; Indels 149; Gaps 26;  
QY 13 DGTSCAGQLVPEANT-AEPISEMPYAGATAATAATAGQVNMIDPIMNNYVAPQGEFTI- 70

Db 178 WSTSETQKILFKQSLGPNLNLPHLAKLYVAMSGSIEVRESISGSGVFGKLAIVVP 237  
Qy 130 PGF-AAQNIQAATMFPHVIAADVRLVLEPIEVLDPVNRVLFHNNNDNAPTMRVLCMLY-- 186  
Db 238 PGIDPVQSTMLQ---YPHVLFDAQVEPVITIPDLRNSLYHLSDDTTTSLVIMIND 294  
Qy 187 --TPRAGSSSGTDPFVIAGRVLCPSDPDSFSLFVLPNPNVQKTKPFSVPNPLNTLSN 244  
Db 295 LNPYANDSNSSGC---IVT--VERKPGDPDFKHLKPPG-----SMLTH 334  
Qy 245 SRVPS 249  
Db 335 GSIPS 339  
RESULT 4  
US-60-427-045-310  
; Sequence 310, Application US/60427045  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Trustees of the University of Arkansas  
; APPLICANT: O'Brien, Timothy  
; APPLICANT: Beard, John  
; APPLICANT: Underwood, Lowell  
; TITLE OF INVENTION: Cal25 Gene and its Use for Diagnostic and Therapeutic  
; TITLE OF INVENTION: Interventions  
; FILE REFERENCE: 022438.44514  
; CURRENT APPLICATION NUMBER: US/60/427,045  
; CURRENT FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 314  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 310  
; LENGTH: 10431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-427-045-310

Query Match 3.8%; Score 109; DB 7; Length 10431;  
Best Local Similarity 18.3%; Pred. No. 34;  
Matches 99; Conservative 81; Mismatches 222; Indels 138; Gaps 22;  
Qy 5 SKDAPTNMDGTSGAGOLVPEANTAEF-ISMPEVAGAAATAAGOVNMDPWHMNNYVQA 63  
Db 8743 SKTEALSGLRTSTPG---PAOSTISPEISLETITRISTPLTTGSAEMTTPKTKHSGAS 8799  
Qy 64 PQGEFTISPNT--PGDILFDLQGLHNPFLSHLAQYNG-----WVGNKVK----- 110  
Db 8800 SGGTFTLDTSSRASPWG---THSAATHRSPHSGMTTPMSRGCPEDVSNPSPRSVEKTSPP 8855  
Qy 111 -VLLAGNAFTAGKII-----ISCIPPPGF- 132  
Db 8856 SSLVLSAVTSPSPLYSTPSESSHSLRVTSLFTPVMMKTTDMLDTSLEPVTTSPSPSMN 8915  
Qy 133 --AAGNISIAQATM-----FPHVIADRVLEPI--EV 160  
Db 8916 ITSDESLATSATMETEAIAQLSENTAVTQMGTISARQEFYSYIGLPPEPSKVTSPVTVSS 8975  
Qy 161 PLEDVRNVLFRHNDNAPTMRVLCMLYTPLRASGSS--SGTDPFVIAGRVLCPSDP 214  
Db 8976 TIKDIVSTIIPASSIETRIEMESTSLTPTRETSTSOEIHSAKPSTVPYKALTSATIE 9035  
Qy 215 FSFLVLP-----PNVEOKTKPFSVPNPLNTLSNRVPSLKSMMVSRDHGQMVQFQNGR 270  
Db 9036 DSMTQVMSSSRGSPDQSTMSQDISSEVITRLTSTPIKAESTEEMITITQTGPGATSRT 9095  
Qy 271 VTLDQO---LOGTFTT-----SASOLCKIRG-----SVFHANGNGYNLTELD 310  
Db 9036 LTLDSTTTFMSTGTHSTASQGFSSHQMTALMSRTPGDVPWLSHPVSEEAASSAS-FSLSD 9151  
Qy 311 GSPYHAFESPAPIGPDLGECDDHWEASPTQFNTGDVTKQINV-----KQESAFAPHLG 365  
Db 9152 -SPVMTSSSPVSSLTLPD-----SIHSSSLPVTSLSLTGSLVKTKTELLGTSSEPTSPPNLS 9206

Qy 366 TIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWVTPRYGSTLTLEAAQLAPPIYPPG 425  
Db 9207 STSAEILATTEVTID--TEKLEMTNVVTSYTHESPSSVLA--DSVTTKATSSNGITYPTG 9263  
RESULT 5  
PCT-US02-33645-26  
; Sequence 26, Application PC/TUS0233645  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of the University of Pennsylvania  
; APPLICANT: Wilson, James M.  
; APPLICANT: Gao, Guangping  
; APPLICANT: Roy, Soumitra  
; TITLE OF INVENTION: Simian Adenovirus Nucleic Acid and Amino Acid Sequences, Vectors  
; FILE REFERENCE: UPN-02677PCT  
; CURRENT APPLICATION NUMBER: PCT/US02/33645  
; CURRENT FILING DATE: 2002-11-20  
; PRIOR APPLICATION NUMBER: US 60/331,951  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/366,798  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 931  
; TYPE: PRT  
; ORGANISM: simian adenovirus SV-1  
PCT-US02-33645-26

Query Match 3.7%; Score 106.5; DB 1; Length 931;  
Best Local Similarity 19.2%; Pred. No. 1.6;  
Matches 105; Conservative 68; Mismatches 226; Indels 147; Gaps 24;  
Qy 9 PTNMDGTSGAGOLVPEAN-----TAEPISMEPVAGAAATAAGOVNMDPWHMNNYVQ 62  
Db 226 PTNKG---GOATPSAQDVQNPQLQFPFASTNVANTPKAVLYAEDVS-----IE 271  
Qy 63 APOGEFTISPNTPG---DILFDLQGLHNPFL-----SHLAQYNGWVGNKVKVLLA 114  
Db 272 ADPTHLVKPTVTEGITSSEALLTQOAPNPNRYTAFRDNFICLMYNTSTGNKV---LA 328  
Qy 115 GNAFTAGKIIISCIPPGFAAQNISIAQATMPHV-----IADRVLEP 157  
Db 329 GOASOLNAV-----DLQDRNTELSYQLMDALGDRSRYFSMMNQAVDSYDPDVRIEN 382  
Qy 158 IEVPLEDRNVLFRHNDNAPTMRVLCMLYTPLRASGSSG--TDPFVIAGRVLCPSDP 215  
Db 383 HGVE-DELPNYCFPLGGMA-----VTDTYSPIKVGGGNGWEANNGVFTTGERGVEIGSN- 435  
Qy 216 SFLFLVPPNVEQKT-KPFSVPNPLN-----TLSNRVPSLKSMMVSRDHGQMVQFQ 267  
Db 436 --MFAMENLQANLWRSFLYSNIGLYLPSDLKITPDNITLP-----ENKNTYQYM 483  
Qy 268 NGRVTLGQLOQTTPTSASQLCKIRGVS---FHANGNGYNLTELDGSPPYHAFESPAP 323  
Db 484 NGRVTPPGLVDTYVNVGARWSPDVDSINPPNHRNAGLRYSMLLGNGRYVPEHIQVQ 543  
Qy 324 GF-----PDGECDDHWEASPTQFNTGDVTKQINVKOESAFAPHLGTIQADGLSD 374  
Db 544 KFFAIKNNLLLPSTYEWNER-----KDVNMLQSSLG---NDRLVDSGAS 587  
Qy 375 VSVNTNMTAKLGWVSPVSDGHRGVDVDPWVTPRYGSTLTLEAAQLAPPIYPPGCEAIVFPM 434  
Db 588 RPSDINLVANF---FPMA-----HNTASTLEAMLRNDTNDQSFENDYLCAAN 630  
Qy 435 SDFPIAHGTNGLSPCTIPOE-----FVTHFNEQAPTGCE-----AALLHYLD 478  
Db 631 MLYPIP--ANATSVISIPSRNNAAFRGSFTRLKTKETPSLGSFGDPFYFVYSGSIPYLD 688  
Qy 479 PDTHRN 484  
Db 689 GTFYLN 694

```

RESULT 6
US-10-258-106-1
; Sequence 1, Application US/10258106
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; TRIBOULEY, Catherine M.
; APPLICANT: KEARNEY, Liam; GRIFFIN, Jennifer A.
; APPLICANT: NGUYEN, Dannel B.; BANDMAN, Olga
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; KHAN, Farran A.
; APPLICANT: WALIA, Narinder K.; YAO, Monique G.
; APPLICANT: PATTERSON, Chandra; BURRILL, John D.
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.
; APPLICANT: RECIPON, Shirley A.; LU, Yan
; APPLICANT: POLICKY, Jennifer L.; THORNTON, Michael B.
; APPLICANT: TANG, Y. TOM; HAFALIA, April J. A.
; APPLICANT: ELLIOTT, Vicki S.; BAUGHN, Mariah R.
; APPLICANT: WALSH, Roderick T.; RAMKUMAR, Jayalaxmi
; APPLICANT: BOROWSKY, Mark L.; AU-YOUNG, Janice K.
; APPLICANT: HILLMAN, Jennifer L.; GURURAJAN, Rajagopal
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0076 USN
; CURRENT APPLICATION NUMBER: US/10/258,106
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/12992
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/199,021
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/200,226
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,339
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/203,505
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/205,654
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/207,739
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/208,795
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2890544CD1
US-10-258-106-1

Query Watch          3.6%; Score 103.5; DB 6; Length 1210;
Best Local Similarity 20.9%; Pred. No. 4.4;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

Qy 9 PTNMDGTSGAGQLVPEANTAEPISEPEVAGAAT-AAATAGOVNMIDPWIMNNVQAPQGE 67
Db 673 PVRMD---NAVPIVQAPAAQPIQIQ--SGVLTOGSCTPLMVATLHPQVA----- 717
Qy 68 FTISPNNTPGDIILFDLQLGPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIIISC 127
Db 718 -TTPQVA---VPFTLSCAAGRALVEQTAALVLAQMPGGTQ-QILLPS---TWQQL---- 765
Qy 128 IPPGFAAONTISIAQTMPFHIVADRVLEPIEVPLEDVRNVLEFHNDNAPTMRVLCMLYT 187
Db 766 --PCVALHN-SVQPTAMPEAMSGQ-----QLADWRNAHSHGNOYSTIMQOQPSLLIN 815
Qy 188 -----PLRAS-----GSSSGTDPEFIAGRV--LTCPSP 213
Db 816 HVTLTATQPLNGVAHVVRQOQSSLLPSKKKQSQAPYSSKSSLD--VLPQSVYSLVGSSP 873

```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10366-64

Query Match      3.4%  Score 99; DB 1; Length 4961;
Best Local Similarity 19.3%  Pred. No. 85;
Matches 129; Conservative 80; Mismatches 254; Indels 204; Gaps 29;

QY 12 MDGTSG----AGQVPEAN-----TAEPISMEFVAGATAAATAGQVANNIDPWNNY 60
Db 2911 LNDRTGWSVASSLSDLNQNFITVTAKDGNPLSQATVHTVTEENHTPEFSQSH 2970
QY 61 VOAPGEGFTISPNTPGDIL-----FDLQGLPHLPFFLSHLAQMYNGW 103
Db 2971 MSA-----TIPESHGISIVRTVSARDRAAMGLIKYSISSGNEEGIF-----AINSS 3019
QY 104 VGNMKYKVLLAGNAFTAGKIIISCTIPGFAAQNISIAQATMFPVHIADRVLEIEVPLE 163
Db 3020 TGIILAKALDYELCKHEMTISAIDGQWART---GYCSVTNVI-DVNDNSPVFLSDD 3075
QY 164 DVNRVLFHNDNAPTMRVLCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPP 223
Db 3076 YEPTVL-----ENAPSGTTV---IHLNATDADSGTNAVI-----AYTVOSD-SDLFVIDP 3122
QY 224 NV-----EQK-----TKPFSVP-----NLPPLNTLSNRVPSLIKSMV 256
Db 3123 NTGVIITQGLDFETKQSYHLTVKAENVPDEERCFSFATVNIQLKG-TNEYVPRFVSKLYY 3181
QY 257 -----SRDHGQMVQFQNGRVTLDGQ----- 277
Db 3182 FEISEAPKGTIVGEFVAFSDRLDGTGEVHYLIFGNSRKKGFQINKGQIIVSGILDPR 3241
QY 278 ---QGTPTSASQCKIRG-----SVFHANGG-----NGYNLTELDGP----- 313
Db 3242 KEERVSUKVLAKNFGSIRGADIDEVTVNVTLDANDPPIFTLNISVQISEGVPIGTHVT 3301
QY 314 -YAFESPAPIGPFDLGECDWHM-EASPTQFNTGDVVIQINVKQE-----SAF 360
Db 3302 FVSAFSDS---IPSWRSRFSYFTGSGNENGAFSINPQTGQITVTAELDRETLPYNLSVL 3358
QY 361 APHLGTTQADGLSDSVSYNTNMIARLGVSPVSDGHRGDVDWPVPIRGYSTLTLEAAQLAPP 420
Db 3359 AVDSGTSATGASALLVLTLEDINDNGPMLTVSEGEVME-----NKRPGTLVMTLQSTD 3412
QY 421 IYPFGGEAIVFFMSDRP-----IAHGTVG-LSVPCITIPQEFVTHFV-----NEQAPTRG 469
Db 3413 DLPNQGPFTYLLSTGPATSYFSLTAGVLTSTREIDREQIADFLSVVTKDGVQPM 3472
QY 470 EAALLHYLDPDTHRLNCEFKLYPEGFTCVPNSSGTGPQTLPIGVFVFSWYSRFRYQLK 529
Db 3473 STGTVHITVIDQDN-----PSQSR-----VEIFVNYGNLFP 3508
QY 530 PVGTAGP 536
Db 3509 ILGSVKP 3515

RESULT 10
US-10-258-951-67
; Sequence 67, Application US/10258951
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murock, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoqing
; APPLICANT: Kabnick, Karen
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Xie, Qing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50025
; CURRENT APPLICATION NUMBER: US/10/258,951
; CURRENT FILING DATE: 2002-10-28

; PRIOR APPLICATION NUMBER: PCT/US01/13360
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/199,963
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/203,336
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/207,087
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/207,546
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-951-67

Query Match      3.4%  Score 98; DB 6; Length 490;
Best Local Similarity 22.8%  Pred. No. 3.5;
Matches 104; Conservative 56; Mismatches 162; Indels 134; Gaps 27;

QY 101 NGWGNMKYKVLLAGNAFTAGKIIISCTIPGFAAQNISIAQATMFPVHIADRVRL 155
Db 51 NGWTASDPVH---GYWFRAGDHVSRNIPVATNNPARAVQEEETDRRF-----HLLGDPQ-N 101
QY 156 EPTEVPLEDVR-----NVLFHNDNAPTMRVLCMLYTPLRASGSSGTDPPF 201
Db 102 KCTLSTRTDRESDAAGTYVFCVERGNKKNYKYDQLSVNVTALTTHMP-----TF 150
QY 202 VIAG-----RVLTCPSPDFSLFLVPPNVEQTKP-----FVSPNLPPLNTLSNRVP 248
Db 151 SIPCTLESHPRLNLTCS-----VPWACEQGTPTITWMCASVSSLD-----P 192
QY 249 SLIKSMVMS-----RDHGQMVQFQNGRVTLDGQLOGTTPITSASQL-----CKIRGSVFH 297
Db 193 TITRSSLMLSLIPQPDHGTSLTQCO---VTLPG--AGVTMTAVRLNLSYPPQNLTTVTEQ 247
QY 298 ANG-----GNGYNLTELDGPYH---AFESPAPIGPFDLGECDW---HMEASPTQFNT 345
Db 248 GDGTASTTLTRNGSALSLEQSLHLVCAVDSNPP-----ARLSWTGSLTSPSSNNL 301
QY 346 GDV-IKQINVKQESAFAPH-----LGTIQADGLSDSVSYNTNMIARLGVSPVSDGHRGDVD 400
Db 302 GYLELPRVHVXKDEGEFTCRANPLGS-QHLSLS-LSLQNEYTKMKRPISGVTILGATGGAG 359
QY 401 P-----WIPRYGSTLTLEAAQLAPPYPPGF--GEAIVFFMSDFPIAHGNTGLS 447
Db 360 ATALVFLYFCIIFVVR--SCRKKSARPAVGVGDTGMEDANAVRGSASQGLPTLESWKD-G 416
QY 448 VPCITIPQEFVTHFVNEQAPTRGEAALLHYLDPDTHR 483
Db 417 NPLKKPPPAV-----APSSGEGEGLHYATLSFHK 445

RESULT 11
PCT-US02-39126-6
; Sequence 6, Application PC/TUS0239126
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: KABLE, Amy E.
; APPLICANT: CHIEN, David
; APPLICANT: WILSON, Amy D.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: GORVAD, Ann E.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: JIN, Pei
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
```



```

; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: KHARE, Reena
; APPLICANT: LEE, Sally
; APPLICANT: HAWKINS, Phillip R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: LEE, Soo yeun
; APPLICANT: SPRAGUE, William W.
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: KINASES AND PHOSPHATASES
; FILE REFERENCE: PF-1315 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/39126
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/340,235
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/343,007
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/343,546
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/354,388
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/357,675
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 1328
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7503214CD1
; PCT-US02-39126-6

Query Match          3.4%; Score 97.5; DB 1; Length 1328;
Best Local Similarity 20.0%; Pred. No. 17;
Matches 108; Conservative 57; Mismatches 207; Indels 169; Gaps 26;

QY 84 QLGPHLNPFLESH---LAQYNGW-VGNMKVKVLLA---GNAFTAGKIISCIPPGF----132
Db 472 KIGVRLNLLGRKSLSEKMNNDVQGFSSVLAHDVGVQAQAERFLKPPVWYLRK 531

QY 133 AQNTISIA---QATMPPHVIADRVLEPTEVLEPDRVNVLFNNDNAPMRL-VCMLYTP 188
Db 532 LVONLLIRRFKKTIEHSROERLNFWDIIFEATNEVT---NGLRFPVLVIEPTKVYQP 589

QY 189 LRASGSSCTDPEVIAGRVLTCPSPDFSEFLVPPNVEOKTKPFSVPNLPLNTLSNRYP 248
Db 590 SYVSINNEAEERTV-----SLHVSPT---EMQMHENFTASSIKGIRFF 632

QY 249 SLTKSMVSRDHGMVQFONGRVTLDDQLQGTTPTSASOLCKIRGSVFH-ANG-----GN 302
Db 633 SLVKEMITNT-----AGSTVELEGTDGDT-----LEYEYDHANGERVVLC 675

QY 303 G-----YNLTELD---GSPYH-----315
Db 676 GTYGIYVAGRDLNSQVRIAIEKIPEDRSYSQLPHEEIALHKYLRNITVQYLGVSSENG 735

QY 316 ---AFESPAPIG-FPDLGECDDHMEASPTQFNTGDVVIQVINVKQESAFAPHLGTIQADG 371
Db 736 YKIFMEQVPGGSLSLALLSKSGPKMKPEIKFYTKOILEGLKYLHENQVHR--DIKGD- 792

QY 372 LSDSVSYNT-----NMIAKLGWVSP-----VSDGHRGDVDVDPWVP 405
Db 793 --NLVNTYSGVVKISDFGTSKRLAGVNCPTETFTGTLQYMAPEIIDQGRGYPGAPADIW 850

QY 406 RYGSTLTEAAQAPPIYPFGCEAIVFMSDPPIAHGTNGLSVPCPTIPOEFVTHFVNEQA 465
Db 851 SLGCTTIEMATSKPPFHELGEPQAAFMFKYGMFKI-H-----PE-----IPEAL 892

QY 466 PTRGEAALLHYLDPDTHRLNGLGEFKLYPEGFMTCVPNSS-----GTCPQ-----TLPLIN 513
Db 893 SAEARAFILSCFEPDPHKKRATTAEELLREGFLROVNGKKNRIAFKPSGPGVVLALPTQ 952
```

```

QY 514 G 514
Db 953 G 953

RESULT 12
US-10-052-648A-39
; Sequence 39, Application US/10052648A
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Foad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 2126
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-052-648A-39

Query Match          3.4%; Score 97.5; DB 6; Length 2126;
Best Local Similarity 20.5%; Pred. No. 33;
Matches 110; Conservative 57; Mismatches 161; Indels 209; Gaps 29;

QY 15 TSGAGOLVPEANTEPISEMPVAGAATAATAGQVNNIDPWIMNNYVQA-----63
Db 1132 TTSTGVVSGGLTTLVSETPTLSSAVSSSTAPAVTVS--TTSQPQVAFSTSGSIASSTG 1189

QY 64 --PQGEFTISPNNTPGDILFDLQGLHLPNPFLLSHLAQMYNGWGNMKVKVLLAGNAFTAG 121
Db 1190 SPPSGTFTTTGTTVSSVAV-----PNAKPPTVLLQQ-----VAGN--TAG 1228
```

----- 489

Qy 332 DWHMEASPTTQNFNDVYKQINVKQSFAFPHLGTIQADGLS 373  
| : : : | : : : |  
nb 490 --HARLDGJLSKMDTEKAASNHVKPK----PELTSLQGQPAS 525

```

RESULT 14
US-09-724-676A-91417
; Sequence 91417, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91417
; LENGTH: 584
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-724-676A-91417

```

Query Match	3.3%	Score 97;	DB 5;	Length 584;
Best Local Similarity	20.9%	Pred. NO. 5.5;		
Matches	84;	Conservative	55;	Mismatches 163;
				Indels 100;
				Gaps 18;

QY 15 TSGACQLYPEANTAEPTISMEPVAGATAATAGQVNMDPWNNNVVQAPEGFTISPNN 74  
:  
Db 181 SSMACKAVAITQTSSSVRLPPAAAEHSPQTAAGMPSVASP--HPDPECKQOITLQP-- 235

Qy	75	TEGDILFDLQGLHNPFLSHL-----ADMYNGWGNM-----KVKVLLAGNAFT-----AG 121
Db	236	TPG-----LPSPTHLEF-SHLPLHSQQSQRTPYNMPVGGIIHWPAGLTYSTFEVPLQAG 288

Qy	122	KLIISCP-----GFAQNTISIAQMFHVIADRVLPLEIEPVEEDRVNVLHNN	34/5
Qy	289	PVLQTL-IPAVSVVHRTLTGTRNTVEGTTNPAGVAELSSVPC-IPICQIRVPGLO-	223
Db			

	QY	174	DNAPMRLVCMIIITPERAGSSGGSGGIEFVPKRVLTLLD       :       :    -----LSTPCLQSPLSMETVINGLANTNNAPQ-----VHPFGALNAVGLQ 3899
	Db	346	-----LSTPCLQSPLSMETVINGLANTNNAPQ-----VHPFGALNAVGLQ 3899
	OY	224	----NVGEQTKTPEFVNPLNTLSNSRVPSPILKSMMVRDHGMVFQNGRTVLDDQLQG 2779

Db 390 VLTNPSSSSPAQAHPTGLQILNIALPTLPSV-----SQVAVDAQAP 435

QY 280 TPTSAQSQCKIRGSVFHANGGNGYNTEL-DGSPYHAFES-----PAPIGPDLGEC 331

436 EMPASQSKACETQPKQTSVASANQSVRTESPOGLFTVORENNAKVLNPPAPAGD-----489

Db 490 --HARLDGLSKMDTEKAASANHVRPK----PELTISIQQGPAS 525

US-09-724-676-91423  
; Sequence 91423, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing

```

; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91423
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-09-724-676-91423

**Qy** 122 KIIISCIP-----PGFAAQ-NISIAQAQMFPHTADRVLEPIEVPLEDRVNVLFHNNDN 175  
|:: | | : : : : | : : | : : |  
**Dd** 1229 VAIVTSVTTFPPAMASQPSPILGSSFSAP-TIAETVV-----VSAHSLDK 1274

```

Qy 176 A---PTWRLVCMLYTPLRASGSSGT-----DPVIAGRVLTCPSPDF3FLFL 220
      |         |         |         |         |         |         |
Db 1275 ASHSSTAGLGLSFCAP--SSSSSGTAVSSVSQFIVHPLVISAATSP-----VL 1325

```

QY	221	VPPNVEQKTKPF-SVENPLNTLSNSRVPSLIKSMMSRDHGQMVFQNGRVILDDGLOG	2719
Db	1326	PQPAVTSTPLLDPVNIPVLQPVANPVAQTLLHSQPALL-----	1370

QY 280 TTPTSASQLKIRGSVFHANGNGYNLTDLSPGSHAFESPAITGFDPDGLGCDWHMEASP 339  
| | | | | : | | | |  
Db 1371 --PNQPHTHC-----PEMDADT---QSKAP-GIDD----- 1394

QY	340	TTQFNIGVGLKQLINVAQSGKPAH	LEIIGGCGGGG	.....	
Db	1395	-----IKTLEELRSLFSEHSSGTOHA	-----SVSLETPLVET	-----VTP	-----G 1433
QV	398	DVDPEWVIPRYGSSILTEAAQLAPPIYPGGEIAIVFMSDFPIAHTNGLSV	-----PCTIQPEF	456	

Db	1434	IPTTAVAPSKLMTSTTSTCLPP-----TNLPL--CTAGMPVMPVGTPOQV	1476
Qy	457	VTFVNEQATRGEAALLHYLDPDTHNLGEFLKYPEGFMTCVPNSSGTG--PQTPL	511

```
DB      1477 ST-----PCTHAS-----APASTATGAKPGTTP 1499
```

\*\*\*\*\*  
RESULT 13

US-09-724-676-91417  
; Sequence 91417, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD

```

; TITLE OF INVENTION: Variants of alternative splicing
;
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222

```

```

; NUMBER OF SEQ ID NOs: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91417
; LENGTH: 584
; TYPE: PRN

```

; ORGANISM: Homo sapiens  
 US-09-724-6776-91417  
 Query Match 3.3% Score 97; DB 5; Length 584;  
 Best local similarity 20.9% pred. No. 5-5;

[illegible][illegible]

DB 236 IPG-----LPFRRHNE SHFENQSQQQKATAMTTCSTT  
QY 122 KIITSCIPP-----GFAQNISIAQATMPPHVIADRVRYLEIPEVLEDVRNVLFHNN 173  
DB 289 PVOLT-IPAVSVYHRTLGPHRRHTVTEVSGTNPAGVAELSSVWPC-IPGCIQIRVPGQLN- 345

Qy 174 DNAPTRLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDSFLVLP-----223  
 ||| : : : : :  
 Db 346 -----LSTPGLQSPSLSMETVNI VGLANTNMAQ-----VHPGLALNAVGLQ 389

QY	224	---NVEQRTKPEVPNPPLNTLSNSRVPSPSLIKSMMSRDHGQWVFQNGFVTLDDGQLQG	279
Db	390	VLTPANSSOSSPAPQAHIIEGLQLINLTALPTLPSV-----SQVAVDAGAP	435

Qy 280 TTTTASQLCKIRGSVFHANGGYNLTTEL-DGSPYHAFES-----PAPIGFPDLGEC 33



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:53:05 ; Search time 21 Seconds  
(without alignments)  
2494.917 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDAPTNMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1974.5	68.2	546	2 B37491	major capsid prote
2	1946	67.2	530	2 B37471	capsid protein - N
3	1223	42.2	542	2 S60616	capsid protein - h
4	1151.5	39.8	539	2 S40111	capsid protein - h
5	316	10.9	2344	1 RWRHR	genome polyprotein
6	315	10.9	576	2 A33982	capsid protein - E
7	311.5	10.8	2344	2 S53399	genome polyprotein
8	307.5	10.6	2344	2 S64740	genome polyprotein
9	300.5	10.4	702	1 A48562	coat protein - San
10	277.5	9.6	668	2 JQ2354	capsid protein - f
11	276	9.5	671	1 VCWFF9	coat protein - fel
12	274	9.5	668	1 VCWFFC	coat protein - fel
13	269	9.3	668	1 VCWFF6	coat protein - fel
14	264	9.1	703	1 C48562	coat protein - San
15	261.5	9.0	668	2 JQ2356	capsid protein - f
16	147	5.1	2206	2 S33822	genome polyprotein
17	146.5	5.1	2205	1 GNNY2W	genome polyprotein
18	145.5	5.0	2207	1 GNNY5P	genome polyprotein
19	145	5.0	2206	1 GNNY4P	genome polyprotein
20	143.5	5.0	2194	1 GNNY7	genome polyprotein
21	141.5	4.9	2207	2 S09553	genome polyprotein
22	139.5	4.8	2206	1 GNNY27	genome polyprotein
23	137.5	4.7	2209	1 GNNY3P	genome polyprotein
24	136.5	4.7	2209	1 GNNY2P	genome polyprotein
25	135.5	4.7	2207	1 GNNY1P	genome polyprotein
26	133.5	4.6	2179	1 GNNY4	genome polyprotein
27	128.5	4.4	613	2 T35828	acetolactate synth
28	125	4.3	3473	1 A46112	genome polyprotein
29	125	4.3	3473	2 S27927	polyprotein - rice

RESULT 1

B37491  
Major capsid protein [similarity] - Southampton virus  
N:Alternate names: orf2 protein  
C:Species: Southampton virus  
C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
C:Accession: B37491  
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Science 259, 516-519, 1993  
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-  
A:Reference number: A37491; MUID:93142023; PMID:8380940  
A:Accession: B37491  
A:Status: not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-546 <LAMS  
A:Cross-references: GB:L07418; MID:g1236787; PIDN:AAA92984.1; PID:g295114  
A:Note: sequence extracted from NCBI backbone (NCBIP:123458)  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: glycoprotein  
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.2%; Score 1974.5; DB 2; Length 546;

Best Local Similarity 66.6%; Pred. No. 6.7e-139;

Matches 367; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

QY 1 MMASKDAPTNMDGTSGAGOLVPEANTAPISMEPVAGATAAATAGQVNMIDPWIMNNY 60

Db 1 MMASKDAPQSGADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNNP 60

QY 61 VOAPQGEFTISPNNTPGDITLFDLQGLPHLNFSLHQAQYNGWGVNMYKVLLAGNAFTA 120

Db 61 VOSQGEFTISPNNTPGDITLFDLQGLPHLNFSLHQAQYNGWGVNMYKVLLAGNAFTA 120

QY 121 GRIITSCIPPGFAAQNISIAQATMFPHTVADRVRLPIEVLDEVRNVLPHNNDNAPTMR 180

Db 121 GRIIVCCVPPGTSSTLTAAQATLPHVITADVRLTLEIEMPLEDRNVLHYHTNDQPTMR 180

QY 181 LVCMLYTPURASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPNNVEBQTKPFSPVPLPLN 240

Db 181 LVCMLYTPURASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPNNVEBQTKPFSPVPLPLN 240

QY 241 TLNSNRVPSLIISMVSRDHQWQVQFQNGRVTLTGQLOGTTPTSASQLCKIRGVSFHANG 300

Db 241 TLNSNRVPSLIISMVSRDHQWQVQFQNGRVTLTGQLOGTTPTSASQLCKIRGVSFHANG 300

QY 301 GNGYNLTEDGSPYHAFESPAPIGPDLCECDWHMEASPT-TQFNTPGDVTKQINVKOE-S 358

Db 301 GNGYNLTEDGSPYHAFESPAPIGPDLCECDWHMEASPT-TQFNTPGDVTKQINVKOE-S 358

QY 299 ARTNLNTEVDGKPPFAFDSPAPVPGPDFGKCDWHMRISKTTPNNTGSGDPMRSVSQTNVQ 358

Db 299 ARTNLNTEVDGKPPFAFDSPAPVPGPDFGKCDWHMRISKTTPNNTGSGDPMRSVSQTNVQ 358

QY 359 AFAPHLGTITQADGLSDSVSYNTNNIAKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQJIA 418

Db 359 AFAPHLGTITQADGLSDSVSYNTNNIAKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQJIA 418

Db 359 GFVPHLGSIQDFEVFNHTPTG-DYICTIEWISQPSPTPGCTDINLMBIPDYGSSLSQAANLA 417  
Qy 419 PPIYPGCEAIVFMSDFPIAHGTNGLS----VPTCTIQEFTVTHFVNEQAQPTGGEAALL 474  
Db 418 PPVPPGCEALVYFVSAP---GPNRSAPNDVPCLLPQOEYITHFVSEQAQTGDAALL 474  
Qy 475 HYLDPDTHRLNGEFLKYPCEGTCVPSNSGCTGQPTLPNGVFEVFSWVSRYQLKPVGTA 534  
Db 475 HYVDPDTHRLNGEFLKYPCEGTCVPSNSGCTGQPTLPNGVFEVFSWVSRYQLKPVGTA 534  
Qy 535 GPA-CRIGLR 544  
Db 535 STARGRLVRR 545  
RESULT 2  
B37471  
capsid protein - Norwalk virus  
C:Species: Norwalk virus  
C:Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C:Accession: B37471  
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A:Title: Sequence and genomic organization of Norwalk virus.  
A:Reference number: A37471; MUID:93303939; PMID:8391187  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: genomic RNA  
A:Residues: 1-530 <JIA>  
A:Cross-references: GB:M87661; NID:g1061313; PIDN:AAB50466.1; PID:g1061313  
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)  
C:Superfamily: human calicivirus capsid protein  
Query Match 67.2%; Score 1946; DB 2; Length 530;  
Best Local Similarity 66.8%; Pred. No. 8.5e-137;  
Matches 365; Conservative 67; Mismatches 96; Indels 18; Gaps 7;  
Qy 1 MMASKADPTNMWDTSGAGOLVPEANTAEPTSMPEVAGATAAATAGOVNMIDPWTMNNY 60  
Db 1 MMASKATSDVSGAGQVLVPEVNASDPLMDPVAGSTAVATAGQVNPIDPWIINF 60  
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120  
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMRVIRMLAGNAFTA 120  
Qy 121 GKIIISCIPPGFAAQNISIAQTAPPHVIADRVLEPIEVPLEDVRNVLFH-NAPTMM 179  
Db 121 GKIIIVSCIPPGFSGSHNTIAQTALFPHVIADRVLDPIEVPLEDVRNVLFHNDNRQOTM 180  
Qy 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSPNPL 239  
Db 181 RLVCMLYTPLRGTGGTG--DSFVAGRVMTCPSPDFNLFVLPVPTVEQKTRPFTLPNPL 238  
Qy 240 NTLNSRVPSLIKSMVSRDHQMVQFONGRVTLDDQLQGTTPTSASOLCKIRGSVFHAN 299  
Db 239 SSLNSRAPLPISSIGISPDNVQSVQFONGRCTLDRVLGVTTPVSLSHVAKIRGT----S 294  
Qy 300 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDHMEASPTTQNTGDKVIQINVKQESA 359  
Db 295 NGTVINLTLDGTPHPFEGPAPIGFDPDLGGCDWHIN---MTQGHSSQTYQVDVDTTPT 351  
Qy 360 FAPHLGTTQADGLSDVSYNTNMIAKLGWSPVSDGHRGDVDPWVTPRYGTSILTEAQLAP 419  
Db 352 FVPHLSIQANGIG----SGNVGVLSWISPPSPHPSGQVLDLKPNGYSSITEATHLAP 407  
Qy 420 PIYPGFGGAIVFFMSDDFPIAHGTNGLSVPCTIQEFTVTHFVNEQAQPTGGEAALLHYLD 479  
Db 408 SVYPPGFGVLVFFMSKMP---GPGAINLPCLLPQIEYLSHLASEQAQPTGGEAALLHYVDP 464  
Qy 480 DTHRNLTGEEKLYPEGFMTCVPSNSGCTGQPTLPNGVFEVFSWVSRYQLKPVGTAGPA-C 538  
Db 465 DTGRNLGEEKKAYPDGLTCTVPGNASSGPPQQLPNGVFEVFSWVSRYQLKPVGTASSARG 524

Qy 539 RLGLRR 544  
Db 525 RLGLRR 530  
RESULT 3  
S60616  
capsid protein - human calicivirus (strain Meiksham)  
C:Species: human calicivirus  
A:Variety: strain Meiksham  
C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
C:Accession: S60616  
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Virus Res. 37, 271-283, 1995  
A:Title: Capsid diversity in small round-structured viruses: molecular characterization  
A:Reference number: S60615; MUID:96136658; PMID:8533462  
A:Accession: S60616  
A:Molecule type: genomic RNA  
A:Residues: 1-542 <GRE>  
A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079  
A:Experimental source: strain Meiksham  
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein  
Query Match 42.2%; Score 1223; DB 2; Length 542;  
Best Local Similarity 47.2%; Pred. No. 4.8e-83;  
Matches 264; Conservative 73; Mismatches 188; Indels 34; Gaps 13;  
Qy 1 MMASKADPTNMWDTSGAGOLVPEANTAEPTSMPEVAGATAAATAGOVNMIDPWTMNNY 60  
Db 1 MKMASNDAASTPDGAAG---LVPESSNN-EYMALEPVAGAAALAPVTGQTNIIIDPWIRANF 56  
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120  
Db 57 VQAPNGEFTVSPRNAPGEVLLNLELGPENLYLAHLARMYNGYAGGVEVQVMLAGNAFTA 116  
Qy 121 GKIIISCIPPGFAAQNISIAQTAPPHVIADRVLEPIEVPLEDVRNVLFH-NNDNAPTMM 179  
Db 117 GKLVFAAVPPHPFVPEVNLSPQKITMPEPHVIDRVLTLEPVLLPLPDVRNSFFHYNOKDDPKM 176  
Qy 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSPNPL 239  
Db 177 RIVAMLYTPLRNSG--SGDDVETVSCRVLTRPSDFDFTYLVPPTVESKTKPFTLPILT 234  
Qy 240 NTLNSRVPSLIKSMVSRDHQMVQFONGRVTLDDQLQGTTPTSASOLCKIRGSV---F 296  
Db 235 GELNSRFPVPTDQMYTSPNEVISVQCQNGRCTLDDGELQGTTLQVSGICAFKGEVTAHL 294  
Qy 297 HANGG-NGYNLTLDGSPYHAFES-PAPIGFPD-----LGECDHMEASPTTQNTG 346  
Db 295 HDNDHLNNVTITNLNGSPDPSEDIPALGVDPFQGRVGVISQDKONAAAGHSEPAKNG 354  
Qy 347 -DVIQINVKQESAFAPHLGTTQADGLSDVSYNTNMIAKLGWSPVSDGHRGDVDPWVIP 405  
Db 355 HDVAVTYTAQ---YTPKLGQIQIGWQDLDLTVNQPVKF---TPVGLNDTEHFNQWVVP 408  
Qy 406 RYGSTUTTEAAQLAPPIYPGFGGAIVFFMSDDFPIAHGTNGLSVPCTIQEFTVTHFVNEQA 465  
Db 409 RYAGALNLTNLNAPSAPVPPGERLFFRSHLPLKGGYGNPAIDCLLPQEWVQHFYQENA 468  
Qy 466 PTRGEAALLHYLDPDTHRNLTGEEKLYPEGFMTCVPSNSGCTGQPTLPNGVFEVFSWVSRE 525  
Db 469 PSSEVALVRYINPDYGRALFEAKLHRAQGMVSSNTS--APVVPANGYPRFSDSWVNOF 526  
Qy 526 YQLKPVGTAGPACRLGIRR 544  
Db 527 YSLAPMGTCN-----GRRR 540  
RESULT 4  
S40111  
capsid protein - human calicivirus (strain Bristol isolate B493)

C:Species: human calicivirus  
A:Variety: strain Bristol isolate B493  
C:Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C:Accession: S40111  
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-seq  
A:Reference number: S40111  
A:Accession: S40111  
A:Molecule type: genomic RNA  
A:Residues: 1-539 <RG>  
A:Cross-references: EMBL:X76716; NID:g436410; PIDN:CAA54134.1; PID:g436411  
A:Experimental source: human enteric calicivirus strain Bristol isolate B493  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match 39.8%; Score 1151.5; DB 2; Length 539;  
Best Local Similarity 44.9%; Pred.No.9.e+78;  
Matches 254; Conservative 78; Mismatches 185; Indels 49; Gaps 16;

QY 1 MMASKADPTNMDDTSGAGQLVPEANTAEIPMEPVAGAATAATAGQVNMDPWINNNY 60  
| | | | : | : | | | : | | | : | | | | | | | | | | | |  
Db 1 MKMASNDANPS-DGS--AANLVEPNV-EVMALPEVVGAIAAPVAGOQNVIDPWIRNF 56  
| | | | : | : | | | : | | | : | | | | | | | | | | | |

QY 61 VQAPQGFTSPNNTPGDILFDLQGLPHLPFETLSHAQMNGWGMKMKVLLAGNAFTA 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 57 VQAPGGEFTSPRNAPGCEILSAPGLPDNLPLYSLSRMYGVAGGEVQVILAGNAFTA 116  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 GKIIISICPGFAAQNISIAQATMFHVIADVRLPIEVLPLEDVNRVLFH-NNDNAPT 179  
| | : | : | : | : | | | | | | | | | | | | | | | | : | : |  
Db 117 GKVFIAAVPNFTPEGUSQCVTMFHIIIVDRQLEPLPLPDVRNNFYHYNQANDSTL 176  
| | : | : | : | : | | | | | | | | | | | | | | | | : | : |

QY 180 RLVCMLTYLRASGGSGTPFVIAGRVLCPSDFSLFLVPNVBOKTKPFSPVNLPL 239  
| : | | | | | | : : | : | | | | | | | | | | | | | | | | : |  
Db 177 KLIALMTYLRA--NNAGDDVTFSRVLRPSDFDFIFLVPTVESRTKPTFPVLTV 234  
| : | | | | | | : : | : | | | | | | | | | | | | | | | | : |

QY 240 NTLSNSRVPLSIKSMVMYSRDHGOMQVQFGRVTLDGOLGTQTPTSASQLCKIRGSVFHAN 299  
| | | | : | : | : | : | | | | | | | | | | | | | | | | : |  
Db "235 EEMSNSRFPILEKLYTGPSAFVQVQPONGRCTTDGVLGCTTOLSAVNICNFRGDVTHIA 294  
| : | | | | | | : : | : | | | | | | | | | | | | | | | | : |

QY 300 GGNCY--NLTELDGSPHARES-PAPIGFFDL-----GECDNHWEASPTTFQNT 345  
| : | : | | | | : | : | | | | | | | | | | | | | | | | | : |  
Db 295 GSHDYTNLASQWNNSVDPTTEEIPAPLGTFDFVGKIQGLLTQTTRADGSTRAHKAT-VST 353  
| : | : | | | | : | : | | | | | | | | | | | | | | | | | : |

QY 346 GDVIKQINVRQESAFAPHLGTIQ--ADGLSDSVSNVTNMIAKLGWSPVSDG--HRGDVDP 401  
| | | | -----HFTEKLSGVFTDTNNDFAQQN--TKFTPVGVGIQGDHHQNEPQQ 401  
| | | | : | : | | | | | | | | | | | | | | | | | | | | : |

QY 402 WVTPRYGSTITEAAQAAPPYPFGFEAIYFFMSDEPIAHGTNGLSVPCPTIPOEFETHFV 461  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | : |  
Db 402 WLLPNYSGRTGHNVHLAPAVALPTFGSQLLFFRSTMPCGCGPYMNLDCILLPQEWLHEF 461  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | : |

QY 462 NEQAPTRGEALLHYLDPDTHRLNGEKLYPEGFEMTCVPMSSGTGPQ--TLTPINGVFVF 519  
| | : | : | | | | | | | | | | | | | | | | | | | | | | | : |  
Db 462 QEAAPAGSDVALLRFVNPMDGRVLFECLKHSKYITV---AHTGPYDLVLPNGYFRFD 517  
| | | : | | | | | | | | | | | | | | | | | | | | | | | : |

QY 520 SWVSRYOLKPVGTGACRGLGIRR 545  
| | | : | | | | | | | | | | | | | | | | | | | | | | | : |  
Db 518 SWNQYFTLAPMGNG-----TGRRA 538  
| | | : | | | | | | | | | | | | | | | | | | | | | | | : |

RESULT 5  
RWWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Meyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557; PMID:1840711  
A:Accession: A41039

Db 125 FGRRLVCAIIPGIGQPCLEVRQ---PPHVVIDARSLPEVTTITMPDLRPEMYHPTGDGP 181  
QY 177 PFMRLVCMLYPLPLRSGSSGTDPF-----VIAGRVLTCPSPDFSEFLVPPNVEQTKP 231  
Db 182 LVPTLWVSYYNNL-----INPEGGTTSAIQVTVETRPSEDFEVLIRAPS--SKTVD 231  
QY 232 FSVPNLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVTLDDGQLQCTTP 282  
Db 232 SVNPWLLTT-----PVLTGAGSDNRWCAPIVGLQPVPGGFTSNRRHWNNGSTYGWSS 285  
QY 283 TSASOLCKIRGSVFHANGGNGVNLDELGSPYHAFESP-----APIGFDLGECDWHMEAS 338  
Db 286 PRFDDIDHSGNVSYPTSATNTIETWYANAGTATTNISNIAPDGFDPDGAIPF-----341  
QY 339 PITQFNTGVDVIQINVKQESAFAPHLGTQADGLSDVSVNTNMIKLGWVSPVSDGHRGD 398  
Db 342 GTT-IPTGAWVFGGQVNNASNGTPIVIGIVQA-----YELGF-----ANGAPSS 383  
QY 399 VDPWVIPRYGSLTEAAQL-APPYPPGFE-----AIVFFMSDFPIANGTNGLSVPCPTIP 453  
Db 384 IRP-----VTTTTCGQLVAKVSIYVAIAQNOSSAGIIFLSKGMVS--TPGVAATYTP 434  
QY 454 Q 454  
Db 435 Q 435  
RESULT 7  
S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate B589)  
C:Species: rabbit hemorrhagic disease virus  
A:Variety: isolate B589  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C:Accession: S55399  
R:Rossi, C.  
Submitted to the EMBL Data Library, May 1995  
A:Reference number: S55399  
A:Accession: S55399  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <ROS>  
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A:Experimental source: isolate B589  
A:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: polyprotein  
Query Match 10.8%; Score 311.5; DB 2; Length 2344;  
Best Local Similarity 27.4%; Pred. No. 2.3e-14;  
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;  
QY 17 GAGQLVPEANTAEPISEMPVAGAAAT-----AAATAG-----QV 49  
Db 1768 GKARTAPOGEAAGATTASVPGTTTDLGDPGVVATTSVVTAESSASIAIATAGIGPPQV 1827  
QY 50 NNIDPWNNVVOAPQGEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGVGNKV 109  
Db 1828 DOQETWRTNFIY---NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGMQF 1884  
QY 110 KVVLAGNAFTAGKIIISCIIPGFAAQNISIA---QATMFPHVIADVRLVLEPIEVLVDVR 166  
Db 1885 RFIVAGSGVFGGRVAAVIPPQ-----IEIGPGLVFRQFPHVVIDARSLPEVITMPDLR 1939  
QY 167 NVLEFH-NDNAPTMRVLVCMLYTPL--RASGSSGTDPPVIAGRVLTCPSPDFSEFLVPP 223  
Db 1940 PNMYHPTGDPGLVPLVLVSVYNNLINPFGSGTS-----AIQVTVETRPSEDFEVMIRAP 1994  
QY 224 NVEQTKPFSVPNPLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVTLDD 274  
Db 1995 S--SKTVDSISPAGLTT-----PVLATGVGNDNRWNGQIVGLQPVPGGFTSNRRHWNLN 2046  
QY 275 QGLOGTTPTTSASOLCKIRGSVFHANGGNGVNLDELGSPYHAFESP-----APIGFP 326  
Db 2047 GSTYGWSSPRFADIDHRRGSASYP-GSNATNVLQFWYANAGS---AVDNPISQVAPDGF 2102

QY 327 DLGECDHMEASPTTQFNTGVDVIQINVKQESAFAPHLGTQADGLSDVSVNTNMIKLG 386  
Db 2103 DMSFVFPNGIPGIPAAAGWVGFAIWNSEN-----SGAPNVTVQA-----YELG 2144  
QY 387 WSPVSDGHRGDVDPWVPIPRYGSTLTLEAAQ-LAPPIYPPGFEA 429  
Db 2145 FAT-----GAPGNLQF-----TTNTSGAQTVAKSIYAVVTGTA 2177  
RESULT 8  
S64740  
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
N:Contains: vp60 protein  
C:Species: rabbit hemorrhagic disease virus  
A:Variety: isolate AST/89  
C:Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
C:Accession: S64740; S46944; S49018; S65012  
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
Submitted to the EMBL Data Library, May 1995  
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by  
A:Reference number: S64740  
A:Accession: S64740  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <CAS>  
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
A:Experimental source: isolate AST/89  
R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra  
submitted to the EMBL Data Library, July 1993  
A:Description: Molecular cloning, sequence and expression of the capsid protein gene  
A:Reference number: S46944  
A:Accession: S46944  
A:Molecule type: genomic RNA  
A:Residues: 1650-2344 <BOG>  
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
A:Experimental source: isolate AST/89  
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.  
Virus Res. 27, 219-228, 1993  
A:Title: The amino terminal sequence of vp60 from rabbit hemorrhagic disease virus su  
A:Reference number: S49018; MUID:93255896; PMID:8488721  
A:Accession: S49018  
A:Molecule type: genomic RNA  
A:Residues: 1650-1796 <PAR>  
A:Cross-references: EMBL:Z24757  
A:Experimental source: isolate AST/89  
A:Accession: S65012  
A:Molecule type: protein  
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881; 1936-1938, 'X', 1940-1941 <PAW>  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: blocked amino end; polyprotein  
Query Match 10.6%; Score 307.5; DB 2; Length 2344;  
Best Local Similarity 26.3%; Pred. No. 4.5e-14;  
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;  
QY 12 MDG---TSSAGQLVPANTAE-----PISMEPVAGAAATAA-----ATAG-----47  
Db 1766 MEKARTAPOGEAAGATTASVPGTTTDCMDPGCVVATTSVVTAESSASIAIATAGIGPPQ 1825  
QY 48 QVNNIDPWNNVVOAPQGEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGVGNM 107  
Db 1826 QVQOETWRTNFIY---NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGM 1882  
QY 108 KVVLAGNAFTAGKIIISCIIPGFAAQNISIA---QATMFPHVIADVRLVLEPIEVLPLED 164  
Db 1883 QRFIVAGSGVFGGRVAAVIPPQ-----IEIGPGLVFRQFPHVVIDARSLPEVITMPD 1937  
QY 165 VRNVLEFH-NDNAPTMRVLVCMLYTPL--RASGSSGTDPPVIAGRVLTCPSPDFSEFLV 221  
Db 1938 LRPNMYHPTGDPGLVPLVLVSVYNNLINPFGSGTS-----AIQVTVETRPSEDFEVMIR 1992  
QY 222 PPNVEQTKPFSVPNPLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVTL 272



```
Db 1993 APS--SKTYVDSISPACLLTT-----PVLTVGVNDNRWNGQIVGLQVPGGFSTCNHRWN 2044
QY 273 LDQLOQTPTTASQLOCKIRGSVFHANGGNYLTEL-----DCSPYHAFESP-----APIG 324
Db 2045 LNSTYGVSPRFGDIGHRGSASYP-GNNATNVLQFWYANAGS---AIDNPISQVAPDG 2100
QY 325 FDLGECDDHMEASPTQNTGQVGIQINVKQESAFAPHLGTLQADCLSDSVSNVTNMIK 384
Db 2101 FDMSEVPFNGPGIPAAGWVGFAIWSN-----SGAPNVTTVQA-----YE 2142
QY 385 LGWVSPVSDGHRGDVDPWIPRYGSTLTAAAO-LAPPYPPGFGAI-----VFEMSDPFI 439
Db 2143 LGFAT-----GAGNLOP-----TTTSGSQTVAKSIYAVVTGTAQNAGLEVMASGVI 2191
QY 440 AHGTNGLSVPCITPOEVTHFVNEQAPTRGEAALLHYLDPDTHRNLGFEKLYPEGFMTCV 499
Db 2192 STPSANAITYTQPDRIVT-----TPGTAAAPVCKNTPIMFASVV 2232
QY 500 PNSS-----CTGQPTLPI 512
Db 2233 RRTGDVGNATAGSANGTQYCTGSQPLPV 2259

RESULT 9
A:Title: coat protein - San Miguel sea lion virus (serotype 1)
A:Accession: A48562
A:Molecule type: capsid protein
A:Residues: 1-702 <SEA>
A:Cross-references: GB:M87481; NID:g3334882; PIDN:AAA16217.1; PID:g3334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.6e-14;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

QY 13 DGTSGAGQLVPEANTAEPTSMPEPVAGATAAATAGOVNMID-PW-IMNMYVQAPOGEFTI 70
Db 154 DGPGGADIVTEEGTVVQOQPVPAQSALTTLAAASTGKTVDCEWTFFSYHTA-----VNM 209
QY 71 SPNNTPGDILFDLQGLPHLNPFLSLHAOMYNGWGMKVKVLLAGNAFTAGKIIISCIPP 130
Db 210 SITEAQGKLFRRALSPENLPLRLHSLSYTSWSGCIDVRFTVSGVGKLAALIVPP 269
QY 131 GFAAQNISTAQATMFPPIADRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLYTPIL- 189
Db 270 GI--BPVESPTMLQPHVLFDARQTEPVITFDIRKTLYHSMDDTDTRLVIMVYNELI 327
QY 190 -----RASSGSSGTDPPFIAGRVLTCPSPDFSLFVPPNVNEQTKPFSVFN--LPLNT 241
Db 328 NPYEASEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLDLIPRNS 377
QY 242 --LSNRVPSLTKSMVSRDHGQWQFQNGR-----VTLD----- 274
Db 378 RHWMGNRWSTIDGEVV-----OPRVOSNRHFDFTSTTGWSTPYPIEVTLEKLDRG 432
QY 275 GOLQGTPTTSASQL-----CKIRGSVFHANGGNYLTEL-----DCSPYHAFESPA 321
Db 433 GOYFKVDTDEKSLVCLPDGWDPTTIPTAMTASNGNYDTVAEYRTNNGTHFKGFYING 492
QY 322 PI-----GPPDLGECDDHMEASPTQNTGQVGIQINVKQESAFAP-----HLGTIOA 369
```

```
Db 493 NUTTKVKGSDNLGET---QOTSRTLTFASVG-----NYKDQNTINPHTHKITSNSLVYDA 543
QY 370 DGLSDSVSVNT---NMIKLGWV---SPVSDGHRGDVDPWIPRYGSGTLTEAAQLA--P 419
Db 544 NVVSAATAKTTTWHSTKMSHLGVLVDESVP-----GSDTKVVRITATLP 587
QY 420 PIYPPGFGAIVFFMSDFPIAHGTNGLSVPCITPOEVTHFVNEQAPTRGEAALLHY-LD 478
Db 588 EAFTNG-GNFPVFETTKIIGH-----FDRAHTKCFNSQVLMTMSOKLAENHYTLP 636
QY 479 PDT 481
Db 637 PDS 639

RESULT 10
QY2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of varia
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09718; NID:g305104; PIDN:AAA16485.1; PID:g305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 277.5; DB 2; Length 668;
Best Local Similarity 23.3%; Pred. No. 1.3e-12;
Matches 124; Conservative 75; Mismatches 178; Indels 155; Gaps 24;

QY 34 EPVAGATAATAAGOVNMIDPMIMNMYVQAPOGEFTI-----SPNNTPGDILFDLQGLP 87
Db 145 EPSAQMSTAADMATGKSVDSW-----EAFESFHTSVNMSTSETGKILFKQSLGP 195
QY 88 HLNPLSLHAOMYNGWGMKVKVLLAGNAFTAGKIIISCIPPFG-AAQNISTAQATMFP 146
Db 196 LLNPYLEHLKSLYVAMSGSVGEVRFSGVFGKLAALVVPVDPVQSTSMLO---YP 252
QY 147 HVIADRVVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLY---TPLRAGSSSGTDPFV 202
Db 253 HVLFDARQVDPVIFSLPDLRSTLYHLMPTDITTSVIMVYNDLINPYANDSNSSGC---I 309
QY 203 IAGRVLTCPSPDFSLFLVPP----- 223
Db 310 VT--VETKPGDPFKFHLLKPPGSMLTGHSVPSDLIPKSSSLWIGNRYWSDITDFVVRPFV 367
QY 224 -----NVEQTKTPESVNL-PLN-TLSNRVPSLIKSMVSRDHGQWQFQNGRVTLD 274
Db 368 FOANRHFNENQETAGSAPRFRPITITISEKSKL--GIGVATDY-----IVP 414
QY 275 GOLQGTPTTSASQLCKIRGSVFHANGGNYLTELDSGYHAFE-SPAPIGPDLGEC- 332
Db 415 GIPDGMPTTIAEDLTTPAGD-YAITSNGNDIT--TGSYDSTEVTKNNTNFKMYICGS 471
QY 333 ----WHMEASPTQF-----NTGDVGIQINVKQESAFAPHLGTLQADGLSD 374
Db 472 LORAWCDKKISNTAFITTAKEGNKIRPSNTIDMTK-IAYQDT-----HVG-----BE 519
QY 375 VSVNTNMIKLGWVSPVSDGHRGDVDP-----WVPRYGTSLTEAAQLAPPYPPGFGAI 430
Db 520 VOTSDALALLGYTGIGEAIGASDRVRVRSVLPEVGAR-----CGNHP 564
QY 431 VEFMSDFPTAHGTNGLSVPCITPOEVTHFVNEQAPTRGEAALLHY-LDPDT 481
Db 565 IFYKSIKILGYVIRSDV-----FNSQILHTRQLSLNHYLLSPDS 605
```

VCWVFC  
coat protein - feline calicivirus (strain Japanese F4)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: B40481  
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yam  
Virology 183, 810-814, 1991  
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
A:Reference number: A40481; MUID:91306470; PMID:1853578  
A:Accession: B40481  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <TOH>  
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 9.5%; Score 274; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. No. 2.3e-12;  
Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMPEVAGATAAATAGQVNMIDPWIMNNYVQAPOGEFTI-----SPNNT 75  
DB 133 PEQGTAVGGVIAEFSQAQMSTAADMASCKSVDSW-----EAFSFTSVNMSSTSET 183  
QY 76 PGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGF-AA 134  
DB 184 QGKILFQSLGFLNLPYLEHLSKLYVAMWSGIEVRESISGSGVFGGKLAIVVPPGVDPV 243  
QY 135 QNISTAQATMPPHYADRVLEPIEVPLEDVRNVLFHNNDNAPTMRVLCMLY----TPLR 190  
DB 244 QSTSLMQ---YPHVLFDARQVEPIFTIPDLRSTLYHVMDSDTTTSLLVIMVYNDLNPYA 300  
QY 191 ASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFPSPVNPPLNTLSNSRVP 249  
DB 301 NDSNSSGC---IVT--VETKPGDPFKFLLKPPG-----SVLTHGSIPS 339

RESULT 13  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline cali  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NEI>  
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 9.3%; Score 269; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. No. 5.4e-12;

VCWVFF9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: genomic RNA  
A:Residues: 1-671 <CAR>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019089; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AA823553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;  
Best Local Similarity 26.4%; Pred. No. 1.6e-12;  
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPOGEFTI- 70  
DB 127 DGSITA----PEQGTAVGGVIAEFSQAQMSTAADMATCKSVDSW-----EAFSFTS 173  
QY 71 -----SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTAGKIII 125  
DB 174 TSVNNSSTSETQCKILFKQSLGFLNLPYLEHLSKLYVAMWSGIEVRESISGSGVFGGKLA 233  
QY 126 SCIPPGF-AAQNISTAQATMPPHYADRVLEPIEVPLEDVRNVLFHNNDNAPTMRVLCM 184  
DB 234 IVPVPGVDVQSTSLMQ---YPHVLFDARQVEPIFTIPDLRSTLYHVMDSDTTTSLLVIM 290  
QY 185 LX-----TPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFPSPVNP 240  
DB 291 VYNDLNPVANDANSSGC---IVT--VETKPGDPFKFLLKPPG-----SMLTHG 335  
QY 241 TLSNSRVPSLIKSMYVRDHGQKVFQNGRVTLDCQLQGTFTPSASCLCKTRGVSFHANG 300  
DB 336 SIPSDLIPKTSLSWIGNYWSDTIDP-----VIRPFVQANR 372  
QY 301 GNGYN-----LTELGGSPYHA-----FESPA-PIGFPPDL---GEC--- 331  
DB 373 HFDFOETAGNSPTRPRPISVTITEONGAKLIGVATDIYVPGIDGWPDDTIPGELIPA 432  
QY 332 -DWHM-----EASPTQFNQGVK 350  
DB 433 GDYAITNGTNDITTATGYDTADIIK 458

RESULT 12

[illegible]

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1974.5	68.2	546	1	COAT_SQV3	Q04542 southampton
2	1148.5	39.7	539	1	COAT_LQDV	P54635 lordsdale v
3	316	10.9	2344	1	POLN_RHDV	P27410 rabbit hemo
4	300.5	10.4	702	1	COAT_SMSV1	P6284 san miguel
5	289	10.0	2208	1	POLN_MANCV	Q69014 manchester
6	276	9.5	671	1	COAT_FCQF9	P27406 feline cali
7	274	9.5	668	1	COAT_FCQV4	P27405 feline cali
8	269	9.3	668	1	COAT_FCQV6	P27404 feline cali
9	264	9.1	703	1	COAT_SMSV4	P36285 san miguel
10	146.5	5.1	2205	1	POLG_POL2W	P23069 polioviru
11	145.5	5.0	2207	1	POLG_POL2L	P06210 polioviru
12	145	5.0	2206	1	POLG_POL3L	P03302 polioviru
13	143.5	5.0	2194	1	POLG_HE701	P32537 human ente
14	139.5	4.8	2206	1	POLG_POL132	P06209 polioviru
15	137.5	4.7	2209	1	POLG_POL15	P03301 polioviru
16	136.5	4.7	2208	1	POLH_POL1M	P03300 p genome po
17	135.5	4.7	2206	1	POLG_POL1M	P03299 p genome po
18	133.5	4.6	2179	1	POLG_HRV14	P03303 human rhino
19	124.5	4.3	855	1	POLG_HRV3	Q82081 human rhino
20	124	4.3	2214	1	POLG_CXA24	P36290 c genome po
21	123.5	4.3	2175	1	POLG_BOVEV	P12915 bovine ente
22	119.5	4.1	2196	1	POLG_EC05N	Q9V1j1 e genome po
23	118.5	4.1	2194	1	POLG_EC30B	Q9Wn78 e genome po
24	117.5	4.1	2185	1	POLG_CXB5P	Q03053 c genome po
25	117	4.0	788	1	BCSB_XANAC	P58933 xanthomonas
26	117	4.0	2183	1	POLG_CXB4E	Q86887 c genome po
27	117	4.0	2183	1	POLG_CXB4J	P08292 c genome po
28	116.5	4.0	2206	1	POLG_CXA21	P22055 c genome po
29	115	4.0	2185	1	POLG_SVDVH	P16604 s genome po
30	115	4.0	2185	1	POLG_SVDVU	P13900 s genome po
31	113.5	3.9	2164	1	POLG_HRV89	P07210 human rhino
32	112.5	3.9	2184	1	POLG_EC01F	Q91734 e genome po
33	112	3.9	2195	1	POLG_EC11G	P29813 e genome po

Query Match	68.2%	Score 1974.5	DB 1:	Length 546
Best Local Similarity	56.6%	Pred. NO. 2.le-143		
Matches 367	Conservative 75	Mismatches 96	Indels 13	Gaps 7
Qy 1	MMWASKDAPNTMDGTCAGAGQLVPEANTAEPLSEMPVAGAATAATAGOVNMIDPWIMNNY 60			
Db				
1	MMWASKDAPQASGASGAGQLVPEVNTADPLMEPVAGPTTAVATAGOVNMIDPWIMNNF 60			
Qy 61	VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLQAQMYNGWGNMKVKVLLAGNAFTA 120			
Db				
61	VQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVRLLAGNAFSA 120			
Qy 121	GKIIISCPGFAAQNTISIAQAATMFPHVIAADVRLVLEPIEVLDPVRNVLFHNNDNAPTMR 180			
Db				
121	GKIIICVCPGFTSSSITIAQATLFPHVIAADVRLVLEPIEPLDVRNVLHTNDNQPTMR 180			

QY	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLPIN	240
QY	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLPIN	240
Db	181	LVCMLYTPLRTGGSGNSDSFVAGRVLTAPSSDSFLVLPPTIEQKTRAFVTPNIPLO	240
QY	241	TLNSRVPSLIKSMVMXRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG	300
QY	241	TLNSRVPSLIKSMVMXRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG	300
Db	241	TLNSRFPSLIQMILSPDASQVQVQNGRCLIDGQLLGTTPATISGQLFVRKGI--NOG	298
QY	301	GNCLNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPT--TQNTGDVTKQINVKOE-S	358
QY	301	GNCLNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPT--TQNTGDVTKQINVKOE-S	358
Db	299	ARTLNLTEDVGGFFWAFDSAPVGPDPFGKCDWHMRISKTPNNNGSGDPMRSVSQTNVQ	358
QY	359	AFAPHLGTTIQADGLSDSVSNWNIKLGWSFVSDGHRGDYDPWVPIRYGSLTTEAAQLA	418
QY	359	AFAPHLGTTIQADGLSDSVSNWNIKLGWSFVSDGHRGDYDPWVPIRYGSLTTEAAQLA	418
Db	359	GEVPHLGSIQFDEVFNHTPG-DYIGTIIWISQPSPTPGTDINLWEIIPDYGSSLSQAANLA	417
QY	419	PIYPGPGCAIVFVMSDFPIAHGTNGLS---VPTIPQEFVTHFYNEQAPTRGEAALL	474
QY	419	PIYPGPGCAIVFVMSDFPIAHGTNGLS---VPTIPQEFVTHFYNEQAPTRGEAALL	474
Db	418	PPVFPFGEGALYFVSAFP--GPNRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALL	474
QY	475	HYLDPDTHRLNGELFKLYPFCMTCPVNSGCTGQPTLPINGVFVSVWVSREFYQLKPYGTA	534
QY	475	HYLDPDTHRLNGELFKLYPFCMTCPVNSGCTGQPTLPINGVFVSVWVSREFYQLKPYGTA	534
Db	475	HYVDPDTHRLNGELFKLYPFCMTCPVNSGCTGQPTLPINGVFVSVWVSREFYQLKPYGTA	534
QY	535	GPA-CRGLGIRR 544	
QY	535	GPA-CRGLGIRR 544	
Db	535	STARGRLGVRR 545	
Db	535	STARGRLGVRR 545	

RESULT 2

COAT\_LORDV STANDARD; PRT; 539 AA.

ID COAT\_LORDV

AC P54635;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Coat protein (Capsid protein).

OS Lordsdale virus (Human enteric calicivirus).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI\_TaxID=82658;

RN [1]

RX MEDLINE=96005060; PubMed=7561776;

RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;

RT "Human enteric Caliciviridae: the complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus".

RT J. Gen. Virol. 76:2349-2355(1995).

CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: X86557; CAA60255.1; .

DR InterPro: IPR004005; Calici.coat.

DR Pfam: PF00915; Calici.coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 539 AA; 58775 MW; 5B5C63E7F2C5FD21 CRC64;

Query Match 39.7%; Score 1148.5; DB 1; Length 539;

Best Local Similarity 44.9%; Pred. No. 2.8e-80;

Matches 254; Conservative 77; Mismatches 186; Indels . 49; Gaps 16;

QY	1	MMWASKDAPTNMDGTSGAGQLVPEANTAEPTSMSEPVAGAAATAATAGQVNMIDPWINNY	60
QY <th>1</th> <th>MMWASKDAPTNMDGTSGAGQLVPEANTAEPTSMSEPVAGAAATAATAGQVNMIDPWINNY</th> <th>60</th>	1	MMWASKDAPTNMDGTSGAGQLVPEANTAEPTSMSEPVAGAAATAATAGQVNMIDPWINNY	60
Db <th>1</th> <th>MMWASKDAPTNMDGTSGAGQLVPEANTAEPTSMSEPVAGAAATAATAGQVNMIDPWINNY</th> <th>60</th>	1	MMWASKDAPTNMDGTSGAGQLVPEANTAEPTSMSEPVAGAAATAATAGQVNMIDPWINNY	60
QY <th>61</th> <th>VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVVLVLAGNAFTA</th> <th>120</th>	61	VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVVLVLAGNAFTA	120
QY <th>61</th> <th>VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVVLVLAGNAFTA</th> <th>120</th>	61	VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVVLVLAGNAFTA	120
Db <th>57</th> <th>VOAPGGEFTVSPNAPGEILWSAPLGPDLPNPLSHLSRMKNGYAGGFEVQVILLAGNAFTA</th> <th>116</th>	57	VOAPGGEFTVSPNAPGEILWSAPLGPDLPNPLSHLSRMKNGYAGGFEVQVILLAGNAFTA	116
QY <th>121</th> <th>GKIIICISPCGFAAQNISIAQAATMPHPHVIADVRVLEPLEVPLEDVRNVLFH--NNDNAPT</th> <th>179</th>	121	GKIIICISPCGFAAQNISIAQAATMPHPHVIADVRVLEPLEVPLEDVRNVLFH--NNDNAPT	179
QY <th>121</th> <th>GKIIICISPCGFAAQNISIAQAATMPHPHVIADVRVLEPLEVPLEDVRNVLFH--NNDNAPT</th> <th>179</th>	121	GKIIICISPCGFAAQNISIAQAATMPHPHVIADVRVLEPLEVPLEDVRNVLFH--NNDNAPT	179
Db <th>117</th> <th>GKVIFAAPVNPFTTEGLSPSQVTPMPEHIIIVRQLEPVLIPDPVRNFFHYHNOANDSTL</th> <th>176</th>	117	GKVIFAAPVNPFTTEGLSPSQVTPMPEHIIIVRQLEPVLIPDPVRNFFHYHNOANDSTL	176
QY <th>180</th> <th>RLVCMLYTPURAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLP</th> <th>239</th>	180	RLVCMLYTPURAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLP	239
QY <th>180</th> <th>RLVCMLYTPURAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLP</th> <th>239</th>	180	RLVCMLYTPURAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBOKTKPFSVPNLP	239
Db <th>177</th> <th>KLIAMLYTPLRA--NNAGDDVFTVSCRVLTRPSDFDFELFVPPTVESRTPFVPLVTV</th> <th>234</th>	177	KLIAMLYTPLRA--NNAGDDVFTVSCRVLTRPSDFDFELFVPPTVESRTPFVPLVTV	234
QY <th>240</th> <th>NTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG</th> <th>299</th>	240	NTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG	299
QY <th>240</th> <th>NTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG</th> <th>299</th>	240	NTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLGQLOGTTPTSASOLCKIRGSVFHANG	299
Db <th>235</th> <th>EEMSNRFPPLKLYTGSPSAFVQVQNGRCTDGVLLGTTQLSVAVNICFRGDDVTHIA</th> <th>294</th>	235	EEMSNRFPPLKLYTGSPSAFVQVQNGRCTDGVLLGTTQLSVAVNICFRGDDVTHIA	294
QY <th>300</th> <th>GGNGY--NLTELGDSPYHAFES--PAPIGPDDL-----GECDDHMEASPTQFNT</th> <th>345</th>	300	GGNGY--NLTELGDSPYHAFES--PAPIGPDDL-----GECDDHMEASPTQFNT	345
QY <th>300</th> <th>GGNGY--NLTELGDSPYHAFES--PAPIGPDDL-----GECDDHMEASPTQFNT</th> <th>345</th>	300	GGNGY--NLTELGDSPYHAFES--PAPIGPDDL-----GECDDHMEASPTQFNT	345
Db <th>295</th> <th>GSHDYTNLNASQNSNVDPTPEEIPAPLGPDPFGVKGITQGLLTQTTTRADGSTRAHKAT--VST</th> <th>353</th>	295	GSHDYTNLNASQNSNVDPTPEEIPAPLGPDPFGVKGITQGLLTQTTTRADGSTRAHKAT--VST	353
QY <th>346</th> <th>GDVIKQINVKOESAFAPHLGTIQ--ADGLSDSVSNVNTMIKLGWSVSPYSDG--HRGDVDP</th> <th>401</th>	346	GDVIKQINVKOESAFAPHLGTIQ--ADGLSDSVSNVNTMIKLGWSVSPYSDG--HRGDVDP	401
QY <th>346</th> <th>GDVIKQINVKOESAFAPHLGTIQ--ADGLSDSVSNVNTMIKLGWSVSPYSDG--HRGDVDP</th> <th>401</th>	346	GDVIKQINVKOESAFAPHLGTIQ--ADGLSDSVSNVNTMIKLGWSVSPYSDG--HRGDVDP	401
Db <th>354</th> <th>GSV-----HFTPLGVSQVFTTDTNNDFOAGN--TKFTPVGVIQDGDHGHQNEPOQ</th> <th>401</th>	354	GSV-----HFTPLGVSQVFTTDTNNDFOAGN--TKFTPVGVIQDGDHGHQNEPOQ	401
QY <th>402</th> <th>WVIPRYGSLTTEAAQLAPPIYPPGFGCAIVFVMSDFPIAHGTNGLSVPTCPQEEVTHFV</th> <th>461</th>	402	WVIPRYGSLTTEAAQLAPPIYPPGFGCAIVFVMSDFPIAHGTNGLSVPTCPQEEVTHFV	461
QY <th>402</th> <th>WVIPRYGSLTTEAAQLAPPIYPPGFGCAIVFVMSDFPIAHGTNGLSVPTCPQEEVTHFV</th> <th>461</th>	402	WVIPRYGSLTTEAAQLAPPIYPPGFGCAIVFVMSDFPIAHGTNGLSVPTCPQEEVTHFV	461
Db <th>402</th> <th>WSLPNVSGRTGHNHVLAPAVAPTFFGQLLFFRSTMPGCGSYPMNMLDCLLPQEWLHFY</th> <th>461</th>	402	WSLPNVSGRTGHNHVLAPAVAPTFFGQLLFFRSTMPGCGSYPMNMLDCLLPQEWLHFY	461
QY <th>462</th> <th>NEQAPTRGEAALLHYLDPDTHRNNGELFKLYPFCMTCPVNSGCTGQ--TLPINGVFV</th> <th>519</th>	462	NEQAPTRGEAALLHYLDPDTHRNNGELFKLYPFCMTCPVNSGCTGQ--TLPINGVFV	519
QY <th>462</th> <th>NEQAPTRGEAALLHYLDPDTHRNNGELFKLYPFCMTCPVNSGCTGQ--TLPINGVFV</th> <th>519</th>	462	NEQAPTRGEAALLHYLDPDTHRNNGELFKLYPFCMTCPVNSGCTGQ--TLPINGVFV	519
Db <th>462</th> <th>QEAAPASQSDVALLRFVNPDPDTGRVLFCKLHKSGYIIV-----AHTGPDVLPNGYFERD</th> <th>517</th> </	462	QEAAPASQSDVALLRFVNPDPDTGRVLFCKLHKSGYIIV-----AHTGPDVLPNGYFERD	517

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC -----
DR EMBL: M67473; AAA47285.1; -.
DR PIR: A41039; RRRWRH.
DR MEROPS: C24.001; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR004004; Calici_pol_hel.
DR InterPro: IPR000317; Endoptase_C24.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR Pfam: PF00915; Calici_coat.1.
DR Pfam: PF03510; Peptidase_C24.1.
DR PRINTS: PR00916; 2CENDOPTASE.
DR PRINTS: PR00918; CALICIVIRUSN.
KW Polypeptide; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN 7508 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 21497 21625 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2344 COAT PROTEIN.
FT NP_BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.9%; Score 316; DB 1; Length 2344;
Best Local Similarity 28.3%; Pred. No. 8.8e-16;
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

QY 12 MDGTSAGQOLVPEANTAEPIS-----MEPVAGAAATAA-----ATAG-----47
DB 1766 MEGKARAAPOGGAAGTATTASPGTTDCMDPGVVATTSVITAENSASIAATAGIGPPQ 1825
QY 48 QVNMIDPWTNNYVQAPQGEFTISNPNPDILDLQLGPHLNPEFLSHLAQYNGVGM 107
DB 1826 QVDQETWRTNFY---NDVFTWSVADAPGSLTYVQHSPPNNPETAQLSQVYAGWAGCM 1882
QY 108 KVKVLLAGNAFTAGKIIISCIIPGFAAQNISTA---QATMEPHVIADVRLPIEVPLED 164
DB 1883 QRFVIVAGSGVFGRLVRAVIPP-----IEIGPLEVQRFPFHVVIDARSLPEVITMPD 1937
QY 165 VNLVFL-NNDNAPTMRMLCYLTP--RASGSSSGTDPFVIAGRVLTCPSPDFSFLEFLV 221
DB 1938 LRPNVYHPTGDPGLVPTLVSYNNLINPFGGTS-----AQVTVETRPSEDFEVMIR 1992
QY 222 PNVQKTKPFVSNPLPLNTLSNSRPSLIKSMVSRDHGQWVQFQ-----NGRVT 272
DB 1993 APS--SKTVDSISPAGLTT-----PVLTVGNDNRWNGQIVGLQVPVPGGFSTCNRHWN 2044
QY 273 LDGQLGTTPTSASOLCKIRGSVFHANGNGYNLTET---DGSPYHAFESP-----APTG 324
DB 2045 LNSTGYGSSPFGDIDHRGSAST-SGSNATNVLQFWYANAGS---AIDNPISQVAPDG 2100
QY 325 FPDLCGCDHMEASPTQFNTGDVIKQINVKQESAFAPHLGTLQADGLSDSVNTNIAK 384
DB 2101 FPDMSFVPPNGGPIPAAGWVGFAIWSN-----SGAPNVTTVQA-----YE 2142
QY 385 LGWSPVSDGHRGVDVDPWPVPRYGTSLTEAAQ-LAPPYPPGFGEA 429
DB 2143 LGFAT----GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 2177
```

## RESULT 4

```
COAT_SMSV1
ID COAT_SMSV1 STANDARD: PRT: 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
```

```
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

CC -----



DR EMBL: M87481; AAA16217.1; -.
DR PIR: A48562; A48562.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.6e-15;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

QY 13 DGTSGAQOLVPEANTAEPISMEPVAGAAATAAATAGVNNID-PW-IMNNYVQAPQGEFTI 70
DB 154 DCPGGADIVTEEGQIVVQOQPVPAOSALTTLAAASGKTVDCEWTFFSYHTA----VNM 209
QY 71 SPNPTPGDILDLQLGPHLNPEFLSHLAQYNGVGMKVKVLLAGNAFTAGKIIISCIIP 130
DB 210 STTEAQGKILFSLRALSPELNPLYLRHISLSYTSWSSGIDVRFVTSVSGVFGGKLAALIVPP 269
QY 131 GFAAONISIAQATMPHVITADVRLPIEVPLEDVRNVLFNHNDNAPTMRMLCYLTP-- 189
DB 270 GI--EPVESPTMLQYPHVLFVDARQTEPVITFDIRKTLYHSMDDDTDTTRLVIMVYNELI 327
QY 190 -----RASGSSSGTDPFVIAGRVLTCPSPDFSFLEFLVPPNVEQKTKPFVSPN--LPLNT 241
DB 328 NPYEQSEPKSSCSIT-----VETRPSSDFTFSLKKPPG--SLLKHGSIPLDLIPRNS 377
QY 242 ---LSNSRPSLTLSKMMVSRDHGQWVQFQNGR-----VTLD-----274
DB 378 RHWGMNRMWSTIDGFVV-----QPRVFQSNRHFDEFTTGTGNSPTPYPIEVLTEKLDRG 432
QY 275 GOLQGTTPTSASOL-----CKIRGSVFHANGNGYNLTET---DGSPYHAFESPA 321
DB 433 GQYFKVTDTEKSLVPLGDPGWPDTTIPTAMTASNGNYDYTVAEYRTITNGTHFKGYIMG 492
QY 322 PT-----GFPDLGECGDHMEASPTQFNTGDVIKQINVKQESAFAP-----HLGTIOA 369
DB 493 NLTKVKGSDNLGET---QOTSRTLFASVG-----NYKQNTINPTHTKITSNSLVVYDA 543
QY 370 DGLSDSVNVT-----NMIKLGWV-----SPVSDGHRGVDVDPWPVPRYGTSLTEAAQLA--P 419
DB 544 NNVSAAATAKTTTWHSTMSHLGYVLVDESVP-----GSDSTKVVRITATLP 587
QY 420 PIYPGFGCEAIYFFMSDFDIANGTNGLSVPCIPIQGFVTHFVNEQAPTGEAALLHY-LID 478
DB 588 EAFENG-GNFPVFFFTNKIQIGH-----FDRAHTKCFNSQVLMTSQKLAENHYTLP 636


```

Thu Jan 16 15:59:23 2003

QY 479 PDT 481  
DB 637 PDS 639

RESULT 5

POLN\_MANCV STANDARD: PRT: 2208 AA.  
AC Q69014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Genome polyprotein [contains: RNA-directed RNA polymerase  
(EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
protein)]; Coat protein].  
DE Manchester virus (Human enteric calicivirus).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
distinct from the Norwalk-like viruses.";  
RL Arch. Virol. 140:1345-1356(1995).  
CC -1- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -1- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
(RNA)(N).  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: X86560; CAA60262.1; -;  
DR MEROPS; C24.UFW; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Endopeptase\_C24.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR PRINTS; PR00915; 2CENDOPEPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR Polyprotein; transferase: RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? 2208 COAT PROTEIN.  
FT NP\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 10.0%; Score 289; DB 1; Length 2208;  
Best Local Similarity 22.9%; Pred. No. 9.4e-14;  
Matches 139; Conservative 82; Mismatches 244; Indels 138; Gaps 25;  
14 GTSAGQLVPEANTAEPISEPVAGAA----TAAATAGQVNMIDPMNMNVQAPQGEFT, 69

DB 1670 GTTG-----PTTSHVVVANPEQNPNGAAORLELATGTA-----IQSNVPEAIRNCEFA 1716  
QY 70 I-----SPNNTPGDILFDLQGLPHLNFSLHQAQMYNGWGNMKVKVLLAGNAFTACKI 123  
DB 1717 VFTFAWDRMPTGTFLGSLHNPINPVYTHSLSCWAGWGSFEVRLISGSGVFAGRI 1776  
QY 124 IISCIPGFAAQNISIAQATMFPVHIAVRLVLEPIEVLPEVLEQVNRNVLPHNNDNA-PTMRLV 182  
DB 1777 IASVIPPQ--VDPSSSTRPCVLPVPHAFVDARITEPVSMFIPDVRAVDYHRMDCAEPTCSLG 1834  
QY 183 CMKYTLPLRASGSSGDPFVIAGRVLTCPSPDFSLFLVPP--NVEQTKTPFSVNPENLPLN 240  
DB 1835 FWYQPLLNPFSTAVSTCWVS--VETKPGDFDCLLRPPQGMENGVSPEGLLPRRLG 1892  
QY 241 TLSNSRVSLIKSMVSRDHQGMVQ--FQNGRVTLGDLQGLTTPTSASOLCKIRGVSFHAN 299  
DB 1893 YSRGNRVGGLVGMILVAEHLKQVNRHNSVTFGWSAPVNPMAA-----EIVTQNAHST 1948  
QY 300 GNGCY-----NLTEL-DGSPYHAFESPAPI-----GFPDLGECWHNEASPTTQ 342  
DB 1949 SRHAWLSIGAQNKGPLPFGIPNHPDSCASTVVGAMDTSLGGRPSTGVC-----GPAIS 2002  
QY 343 F-NTGQDVIKQINVKQESAFAPH-----LGTIADGLSDVSV-----N 378  
DB 2003 FQNGNDVVE--NDTPSVMFATYDPLTSGTGVALTNSINPASLALVRISNNDFTSGFAND 2060  
QY 379 TMTAKLGWSPVSDGH--RGVDVDPWVTPRYGSLTTEAAQLA-----PPIYP 424  
DB 2061 KNVVVQSWEMYTQNTQIRGQVTPMSGNTYFTTGTANTLVQERMLSYDGHQAIIYSS 2120  
QY 425 GFGEAIVFMSDFPIAHGTNGLSVPTIPOEFVTHFVNEQAPTRGEAALLHYLDPTHNRN 484  
DB 2121 QLERTAEYFQNDI-----VNTPNSMAVF-----NVTNSA 2151  
QY 485 LGEFKLYPEGMTCPVNSSGTGPOTPLNGVVFVMSVMSRYQLKPVGTA--GPACRLGI 542  
DB 2152 SPQIGIRPDGYMV-----TGGSGVNVPLEPETRF-----QVVGILPLSALSGSGNMR 2202  
QY 543 RR 544  
DB 2203 AR 2204

RESULT 6

COAT\_FCVF9 STANDARD: PRT: 671 AA.  
AC P27406;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Coat protein (Capsid protein).  
GN Cpp76.  
OS Feline calicivirus (strain F9) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OX NCBI\_TaxID=11981;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=92410623; PubMed=1529544;  
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,  
Turner P.C.;  
RT "The complete nucleotide sequence of a feline calicivirus.";  
RT Virology 190:443-448(1992).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=92339673; PubMed=1633955;  
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
RT "Cloning and sequence determination of the feline calicivirus strain  
F9.";  
RT Biochem. Soc. Trans. 20:26S-26S(1992).  
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M86379; AAA79327.1; -  
 DR EMBL: Z11536; CAA77636.1; -  
 DR PIR: B43382; VCMWFC.  
 DR PIR: C45538; C45538.  
 DR InterPro: IPR004005; Calici\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;

Query Match 9.5%; Score 276; DB 1; Length 671;  
 Best Local Similarity 26.4%; Pred. No. 1.8e-13;  
 Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DFTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70  
 DB 127 DGSITA---PEQGTWGVGVIAEPSAQMTAADMATGKSVDSW-----EAFESFH 173  
 QY 71 -----SPNWTGDIPLDGLPHLPFLSHLAQYNGWGNMKVKVLLAGNAFTAGKIII 125  
 DB 174 TSVNWTSETQGRILFKQSLGPNLPLEHLAKLYVWMSGIEVRSISGSGVFGKLA 233  
 QY 126 SCIPPCF-AAQNTSIAQATMFHIVADVRLPIEVLPLEDRVNLPHNDNAPTMLVCM 184  
 DB 234 IVVPGVDVQVSTMLQ---YHVLFDARQVEPIITDRLSTLYHMSDITDTTSLVIM 290  
 QY 185 LY-----TPLRASGSGTDFVIAGRVLCPCPDFSFLFLVPPNVQKTKPFSVNPPLN 240  
 DB 291 VYNDLINPVANDANSNGC---IVT--VETKPGDPFKHLKPPG-----SMLTHG 335  
 QY -241 TLSNRPVSLIKSMVSRDHQWQVQNGRVLDGLOGTTPTSASOLCKIRGSVPHANG 300  
 DB 336 SIPSDLIPKTSLLWGNRYNSDITDF-----VIRPFVQANR 372  
 QY 301 GNGYN-----LTLEDGSPYHA-----PESPA-PIGFPLD---GEC--- 331  
 DB 373 HFDNQETAGWTPRPRIISVITEONGAKLGIGVATDVIQIPDGPWDTTIPGELIPA 432  
 QY 332 -DWHM-----EASPTQFNTGDKV 350  
 DB 433 GDVAITNGTGNDITATGYDTADIIK 458

RESULT 7  
 COAT\_FCVC6  
 ID COAT\_FCVC6 STANDARD; PRT; 668 AA.  
 AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 RA Miyamura K., Yamazaki S., Mikami T.;  
 RT "Sequence analysis of the 3'-end of feline calicivirus genome."

RL Virology 183:810-814(1991).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: D90357; BAA14371.1; -  
 DR PIR: B40481; VCMWFC.  
 DR InterPro: IPR004005; Calici\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85804E503 CRC64;

Query Match 9.5%; Score 274; DB 1; Length 668;  
 Best Local Similarity 31.0%; Pred. No. 2.6e-13;  
 Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI-----SPNWT 75  
 DB 133 PEOGTAVGGVIAEPSAQMTAADMATGKSVDSW-----EAFSPHTSYNWTSET 183  
 QY 76 PGDILFDLGLPHLPFLSHLAQYNGWGNMKVKVLLAGNAFTAGKIIISCIIPPG-AA 134  
 DB 184 QGKILFKQSLGPNLPLEHLKLYVWMSGIEVRSISGSGVFGKLAIVVPGVDPV 243  
 QY 135 QNTSIAQATMFHIVADVRLPIEVLPLEDRVNLPHNDNAPTMLVCMXY----TPLR 190  
 DB 244 QSTSMQLQ---YHVLFDARQVEPIITDRLSTLYHMSDITDTTSLVIMVNDLINPYA 300  
 QY 191 ASGSSSGTDFVIAGRVLCPCPDFSFLFLVPPNVQKTKPFSVNPPLNTLSNRVPS 249  
 DB 301 NDSNSSG---IVT--VETKPGDPFKHLKPPG-----SVLTHGSIPS 339

RESULT 8  
 COAT\_FCVC6  
 ID COAT\_FCVC6 STANDARD; PRT; 668 AA.  
 AC P27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11979;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus."  
 RL J. Virol. 65:5440-5447(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Neill J.D.;  
 RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -





Thu Jan 16 15:59:23 2003

```

DR InterPro: IPR000199; Cys-protease-3C.
DR InterPro: IPR003138; Pico_P1A.
DR InterPro: IPR000081; Pico_P2A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001676; RN.
DR Pfam: PF00073; rhv; 3.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF00680; RNA_Gep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR ProDom: PD001125; Cys-protease-3C; 1.
DR ProDom: PD001274; Pico_P2B; 1.
DR ProDom: PD001306; Pico_P2A; 1.
DR ProDom: PD001306; Pico_P2B; 1.
KW Polypeptide; Coat protein; Core protein; Core protein; Hydrolase; Thiol protease; Myristate.
KW RNA-directed RNA polymerase; RNA-directed RNA polymerase; RNA-directed RNA polymerase; RNA-directed RNA polymerase.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 579 879
FT CHAIN 880 1028
FT CHAIN 1029 1125
FT CHAIN 1126 1454
FT CHAIN 1455 1541
FT CHAIN 1542 1563
FT CHAIN 1564 1746
FT CHAIN 1747 2207
FT LIPID 2 2
FT ACT SITE 1710 1710
FT ACT SITE 1724 1724
FT ACT SITE 1724 1724
FT ACT SITE 1724 1724
SQ SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;

Query Match 5.08; Score 145.5; DB 1; Length 2207;
Best Local Similarity 18.88; Pred. No. 0.0087;
Matches 106; Conservative 72; Mismatches 193; Indels 193; Gaps 24;

QY 22 VPEANTAPISMEPVAGATAAATAGQVNMIDPWIMNMYVQAFQGEFTISP-----NTP 76
DB 365 IPEFDTTPID---IPGEVRNMELAEIDTMIPLNLTNQRKNTMDYRVELINDAAHSDDT 421
QY 77 GDILFDLQGLPHLNFSLH-----LAQWNCWGVNMKVLLAGNATAGKIIISCPGP 131
DB 422 ---ILCLSLSPASDRLAHTMLGEILNYTHWAGSLKFTFLFCGSMATGKLVSYAPPG 478
QY 132 FAQNISIAQATMFPVIAADVRLPIEVPLEDVNRNVLHN--NDNAP-----TMR 180
DB 479 AEAPK-SRKEAMLGTHVINDICLQSSCTMVVPWISNTTYROTINDSFTEGGYISMFYQTR 537
QY 181 LVCMLYTPLRASGSSGTPPFVIAGRVLTCPSPDFSELF-----LVP 222
DB 538 VVVLSTPRKMD-----ILGFVSAC--NDFSRLRLDTHTHSQEAMPQGLDLIE 585
QY 223 PVNEQTKTFSPVNPPLNTLSNR-----VPS-----LIKSMVMSRDHGQ 262
DB 586 GVVEGVTNRALPTLTPANNLPDQSSGPAHSKETALTAVETGATNPLVPSDTVQTRH-- 643
QY 263 MVQ-----PQNGRVTLDDLOQGTTPTS-ASOLCKTRGSVF-----296
DB 644 VIQKTRSESTVESFFARGCAVAIEVDNDAPTCKRASKLFSVWKITYKDTVOLRRKLEFF 703
QY 297 -----HANGCNYNLTELDGSPYHAFESPAPIGPDLCECDWHMEASPT 340
DB 704 TYSREDFMEFTVVTSTYTDANNHALNQVQIMYPGPAPIGKKNYD---TWQTSNPS 760
QY 341 TQNTGVDVQINVKQESAFAPHLGHTQA-----DGLSDVSVNNTNNAKLGWSPVSDGH 395
DB 761 VFYTYGAPPARISV-----PVVGIANAYSHEFDGFAKV-----793
QY 396 RGVDPVWVIFRYSTLTAAQLAPPYPPGFGGAIYVFFMSDFPIAHGTNGLSVPCIIPOE 455

```

```

Db 794 -----PLAQASTEGDSL-----YGA-----SLNDF-----GSLAV----- 820
QY 456 FVTHFVNEQAPTRGEAALLHYLDP 479
Db 821 ---RVVNDHNPTKLTISKIRVYMKP 841

RESULT 12
POLG_POL3L STANDARD; PRT: 2206 AA.
AC POLG_POL3L; 084783; 084784; 084785; 084786; 084787; 084788; 084789;
AC 084790; 098592; 098593; 098594;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: coat proteins vp1 TO vp4; Core proteins
DE P2A TO P2C, P3A; Genome-linked protein vpg; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12088;
RN [1]
SEQUENCE FROM N.A.
STRAIN-P3/Leon/37;
MEDLINE-84170338; PubMed-6324200;
RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,
RA Schild G.C., Almond J.W.;
RA "Comparison of the complete nucleotide sequences of the genomes of
RA the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin
RA vaccine derivative P3/Leon 12alb.";
RT proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).
RN [2]
SEQUENCE FROM N.A.
STRAIN-P3/Leon 12A[1]B;
RX MEDLINE-83299239; PubMed-6310508;
RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,
RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;
RA "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison
RA with poliovirus type 1.";
RL Nucleic Acids Res. 11:5629-5643(1983).
RN [3]
X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.
RP MEDLINE-95120467; PubMed-7820548;
RA Grant R.A., Hirenath C.N., Filman D.J., Syed R., Andries K.,
RA Hogle J.M.;
RA "Structures of poliovirus complexes with anti-viral drugs:
RA implications for viral stability and drug design.";
RL Curr. Biol. 4:784-797(1994).
RN [4]
X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.
RP Hirenath C.N., Grant R.A., Filman D.J., Hogle J.M.;
RA "Binding of the antiviral drug win51711 to the Sabin strain of type-3
RA poliovirus - structural comparison with drug-binding in rhinovirus-
RA 14.";
RL Acta Crystallogr. D 51:473-489(1995).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate " N diphosphate +
CC [RNA](N). THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS
CC SHOWN.
CC -1- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE
CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

```



Query Match	Score	DB 1	Length	DB 2	Length	DB 3	Length	DB 4	Length
Best Local Similarity	19.4%	Pred. No.	0.012						
Matches	94	Conservative	57	Mismatches	166	Indels	167	Gaps	17
QY	70	ISPNPTGDIILF	5.0%	Score 143.5	DB 1	Length 2194			
DB	390	ISAGSDMDLLFNPLDIEGLPLRLNTLGNISRYTHNSGSLEMTFMCGSFMTTGKLI							
QY	125	ISCTPPGFAAGNISIAQATMPHHVADRVLEPIEPLVDVRLVFNHNDNAPTMLVCM							
DB	450	ICYTPPGGSSPTDRM-QAMLAATHVWVDFGLQSSITIIIPWISGSHYR							
QY	185	LYTPLRASGSSGDPFVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFPVNPPLNTLSN							
DB	497	FNTDAKALNANG							
QY	245	SRVPSLIKSMVMVRDHGQWQFQNGRVTLGQLOGTTPTTSASOLCKIRGSVFHANGNGY							
DB	520								
QY	305	NITELDGSPYHAFESPAPIGPDLGECDDHMEASPTTQNTGDVQIKOVNKQESAFAPHL							
DB	542	NLRLRDS							
QY	365	GTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWVTPRYGSTTTEAAQALPPIYP							
DB	587	GVIPS							
QY	425	GRGEAIV							
DB	619	GTAECLEVNFLGRSALVCMRSFEYKNHSTST							
QY	474	LHYLDPTDTHNLGEFLKXPEGFMTCVNPSSGTPQTLPIGVFVSVWRFYOLKPVGT							
DB	676	FYTLREDT							
QY	534	AGPA							

```
FT CHAIN 579 878 COAT PROTEIN VP1.
FT CHAIN 879 1027 CORE PROTEIN P2A.
FT CHAIN 1028 1124 CORE PROTEIN P2B.
FT CHAIN 1125 1453 CORE PROTEIN P2C.
FT CHAIN 1454 1540 CORE PROTEIN P3A.
FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1563 1745 PICORNAIN 3C.
FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;

Query Match 4.8%; Score 139.5; DB 1; Length 2206;
Best Local Similarity 20.8%; Pred. No. 0.025;
Matches 96; Conservative 64; Mismatches 182; Indels 119; Gaps 21;

QY 22 VPEANTAEPISEMPVAGATAAATAGOVNMIDPWIMNNYVQAPQGEFTI---SPNTPGD 78
DB 365 IPEFDVTPID---IPGEVKNMELAEIDTMIPLNLENTKRTMDMYRVRLSDSANLSGP 421
QY 79 ILFDLQLGPHLPFLSH-----LAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGFA 133
DB 422 ILC-LSLSPAADPRLSHTMLGEVLYNYTHWAGSLKETFLFCGMMATGKLLVAYAPPG-A 479
QY 134 AQNISIAQATMPHHVIADVRVLEPIEVLPLEYVRLFH--NNDNAPTRMLVCMLY----- 186
DB 480 QPPTSKEAMLTGTHVLDLGLQSSCTMVVPWISNVTYRQTQDSFTGEGYISMFYQTRIV 539
QY 187 TPLRAGSSSGTDPFVIAGRVLTCTSPDPSFLF-----VPPNVE-----QKT 229
DB 540 VPL-----STPKAMDMLGFSVAC--NDFSRLLRDTHHISOAAMPQGVDDLLITEVAQNA 591
QY 230 KPFSVP-----NLP-----LNTLS-----NSRVPS 249
DB 592 LALSPLKPOSNLPDTHKASPAHSKEVPTLTAVETGATNPLVSDTVQVRHVITQQRSSS 651
QY 250 LKSM-----MYSRDHQ---MVQ---FQNGRVT-----LDQGLQTTPTSASQLCK 290
DB 652 TIESFARGACVAITEVDNEQATNVQKLFATWRITKDTVQLRRKLEFTYVSREFDMEFT 711
QY 291 IRGSVFHANGNGYNLTLDGSPYHAFESPAPGPDLCGDWHMEASPTQFNGDVVK 350
DB 712 FVVTANTNSNGHALNQVQIYIMYPGAPTCKSDWDY---TWQTSNPSIFITYGAAPA 768
QY 351 QINVAQESAFAPHLGTIOA-----DGLSDVSVNTNNIAKLG 386
DB 769 RISV-----PYVGLNAYSHEFDGFAKVPKLSKANDQVG 802

RESULT 15
POLG_POLIS STANDARD: PRT: 2209 AA
AC Q03301; Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;
AC Q84888; Q84889; Q84890;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 01, Last sequence update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)]
OS Poliovirus type 1 (strain Sabin).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299876; PubMed=6310545;
RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,
RA Genba Y., Nakano Y., Imura N.;
RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1
strain genome.";
```

```
RL PROC. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM
CC THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION
CC PROCESS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01150; CAA24465.1; -.
CC PIR; A03899; GNNV3P.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC InterPro; IPR000199; Cys-protease-3C.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR000081; Pico_P2A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR001676; Rhv.
CC Pfam; PF000073; rhv; 3.
CC Pfam; PF005548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC Polyprotein; Coat protein; Core protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
KW CHAIN 2 69
FT CHAIN 70 341 COAT PROTEIN VP4.
FT CHAIN 342 579 COAT PROTEIN VP3.
FT CHAIN 580 881 COAT PROTEIN VP1.
FT CHAIN 882 1030 CORE PROTEIN P2A.
FT CHAIN 1031 1127 CORE PROTEIN P2B.
FT CHAIN 1128 1456 CORE PROTEIN P2C.
FT CHAIN 1457 1543 CORE PROTEIN P3A.
FT CHAIN 1544 1565 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1566 1747 PICORNAIN 3C.
FT CHAIN 1748 2209 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE.
FT ACT_SITE 1712 1712 PROTEASE (POTENTIAL).
FT ACT_SITE 1726 1726 PROTEASE (POTENTIAL).
SQ SEQUENCE 2209 AA; 246576 MW; 9EC1FF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;
Best Local Similarity 19.3%; Pred. No. 0.036;
Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPEANTAEPISEMPVAGATAAATAGOVNMIDPWIMNNYVQAPQGEFTISPNNTP--GDI 79
DB 366 LPEFDVTPID---IPGEVKNMELAEIDTMIPLFDLSAKKKNMTMEMYRVRLSDKPHITDPP 422
QY 80 LFDLQLGPHLPFLSH-----LAQMYNGWGNMKVKVLLAGNAFTAGKTIISCIIPGFAA 134
```

Search completed: January 16, 2003, 09:55:17  
Job time : 23 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:52:25 ; Search time 38 Seconds  
(without alignments)  
2955.149 Million cell updates/sec

Title: us-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPNTNMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12	Q8VA02
2	2846	98.3	545	12	Q9WI82
3	2846	98.3	545	12	Q918C8
4	2840	98.1	545	12	Q918C2
5	2798	96.6	545	12	Q91H12
6	2744	94.8	545	12	Q9IV45
7	2605.5	90.0	544	12	Q66418
8	2124	73.3	539	12	Q9IV48
9	1979	68.3	544	12	Q9DU46
10	1975	68.2	544	12	Q91I85
11	1974	68.2	544	12	Q9QTE7
12	1972.5	68.1	540	12	Q9YS14
13	1968	68.0	544	12	Q918C5
14	1964.5	67.8	546	12	Q9IV47
15	1956.5	67.6	543	12	Q9IV43
16	1956.5	67.6	544	12	Q8VA27

17	1954	67.5	544	12	Q918D1	Q918d1	norwalk-lik
18	1952	67.4	544	12	Q91E32	Q91e32	human calic
19	1946.5	67.2	543	12	Q918B9	Q918b9	norwalk-lik
20	1946	67.2	530	12	Q83884	Q83884	norwalk-lik
21	1942	67.1	544	12	Q9IV36	Q9iv36	human calic
22	1941	67.0	544	12	Q9IV41	Q9iv41	human calic
23	1938	66.9	530	12	Q9Q739	Q9qt39	norwalk-lik
24	1930.5	66.7	531	12	Q8V7J5	Q8v7j5	norwalk-lik
25	1908.5	65.9	517	12	Q9JH41	Q9jh41	norwalk-lik
26	1900	65.6	530	12	Q83876	Q83876	human calic
27	1886	65.1	541	12	Q9IV42	Q9iv42	human calic
28	1274	44.0	522	12	Q8V629	Q8v629	norwalk-lik
29	1268	43.8	522	12	Q8V628	Q8v628	norwalk-lik
30	1260.5	43.5	519	12	Q9YQ22	Q9yq22	bovine cal
31	1228	42.4	540	12	Q917X1	Q917x1	norwalk-lik
32	1223	42.2	542	12	Q66293	Q66293	calicivirus
33	1222	42.2	542	12	Q913B6	Q913b6	human calic
34	1219	42.1	540	12	Q91V50	Q91v50	human calic
35	1215.5	42.0	542	12	Q12305	Q12305	snow mounta
36	1213	41.9	542	12	Q96877	Q96877	snow mounta
37	1211	41.8	538	12	Q91I65	Q91i65	human calic
38	1203	41.5	540	12	Q917W8	Q917w8	norwalk-lik
39	1202.5	41.5	535	12	Q915C6	Q915c6	human calic
40	1202.5	41.5	540	12	Q918B0	Q918b0	norwalk-lik
41	1198.5	41.4	548	12	Q88291	Q88291	small round
42	1196	41.3	537	12	Q913B7	Q913b7	human calic
43	1189	41.1	537	12	Q91H09	Q91h09	human calic
44	1188.5	41.0	548	12	Q917Z5	Q917z5	norwalk-lik
45	1188.5	41.0	548	12	Q916E5	Q916e5	human calic

#### ALIGNMENTS

#### RESULT 1

Q8VA02 PRELIMINARY; PRT: 545 AA.  
AC Q8VA02;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Potsdam 196/2000/DR.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.  
OX NCBI\_TaxID=150135;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/POTSDAM 196/2000/DE;  
RA Kuenkel U., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF439267; AAL32455.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 545 AA; 58601 MW; 7D47D4D9F3521D89 CRC64;

Query Match 99.2%; Score 2873; DB 12; Length 545;  
Best Local Similarity 99.1%; Pred. No. 1e-222;  
Matches 540; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISMPEVAGAAATAAGQVNMIDPWIMNNY 60  
|||||  
Db 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISMPEVAGAAATAAGQVNMIDPWIMNNY 60  
|||||  
Qy 61 VOAPGEFTISPNNTPGDITLFDLQLGPHLNPFLSHLAQMYNGWVGKMKVKVLLAGNAFTA 120  
|||||  
Db 61 VOAPGEFTISPNNTPGDITLFDLQLGPHLNPFLSHLAQMYNGWVGKMKVKVLLAGNAFTA 120  
|||||  
Qy 121 GKIIISCIPPGFAAQNISIAQATMFPHPVIADRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180  
|||||  
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHPVIADRVLEPIEVPLEDVNRVLFHNNDNAPTMR 180  
|||||

QY	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDSEFLVPPNVEQTKPFVSNPLPLN	240		121	GKIIISCIPIPGFAAQNISIAQATMPEPHVIADRVRLPEIETVPLEDRVNVLFHNDNTPTMR	180	
Db	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDSEFLVPPNVEQTKPFVSNPLPLN	240		181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDSEFLVPPNVEQTKPFVSNPLPLN	240	
QY	241	TLNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300		181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDSEFLVPPNVEQTKPFVSNPLPLN	240	
Db	241	TLNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300		241	TLNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300	
QY	301	GNGYNLTLDGSPYHAFESPAPIGFPDLGECGDHMEASPTTQFNTGDIKQINVKQESAF	360		241	TLNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300	
Db	301	GNGYNLTLDGSPYHAFESPAPIGFPDLGECGDHMEASPTTQFNTGDIKQINVKQESAF	360		301	GNGYNLTLDGSPYHAFESPAPIGFPDLGECGDHMEASPTTQFNTGDIKQINVKQESAF	360	
QY	361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPP	420		301	GNGYNLTLDGSPYHAFESPAPIGFPDLGECGDHMEASPTTQFNTGDIKQINVKQESAF	360	
Db	361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPP	420		361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPP	420	
QY	421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOQEFVTHFVNEQAPTRGEAALLHYLDPD	480		361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPP	420	
Db	421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOQEFVTHFVNEQAPTRGEAALLHYLDPD	480		421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOQEFVTHFVNEQAPTRGEAALLHYLDPD	480	
QY	481	THRNLGGEFKLYPEGFMTCVPNSSSGTGPTLPINGVFVSWVSRYQLKPVGTAGPACRL	540		421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOQEFVTHFVNEQAPTRGEAALLHYLDPD	480	
Db	481	THRNLGGEFKLYPEGFMTCVPNSSSGTGPTLPINGVFVSWVSRYQLKPVGTAGPACRL	540		481	THRNLGGEFKLYPEGFMTCVPNSSSGTGPTLPINGVFVSWVSRYQLKPVGTAGPACRL	540	
QY	541	GIRRS 545			481	THRNLGGEFKLYPEGFMTCVPNSSSGTGPTLPINGVFVSWVSRYQLKPVGTAGPACRL	540	
Db	541	GIRRS 545			541	GIRRS 545		
RESULT 3								
ID	Q918C8	PRELIMINARY;	PRT;	545 AA.				
AC	Q918C8;							
DT	01-DEC-2001 (TrEMBLrel. 19, Created)							
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)							
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)							
DE	Capsid protein.							
GN	ORF2.							
OS	Norwalk-like virus NLV/Honolulu/219/1992/US.							
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;							
OC	Norwalk-like viruses.							
OX	NCBI_TaxID=171835;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=HU/NLV/HONOLULU/219/1992/US;							
RX	MEDLINE=97193806; PubMed-9041391;							
RA	Ando T., Monroe S.S., Noel J.S., Glass R.I.;							
RT	"A one-tube method of reverse transcription-PCR to efficiently amplify							
RT	a 3-kilobase region from the RNA polymerase gene to the poly(A) tail							
RL	of small round-structured viruses (Norwalk-like viruses).";							
RL	J. Clin. Microbiol. 35:570-577(1997).							
RN	[2]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=HU/NLV/HONOLULU/219/1992/US;							
RX	MEDLINE=20266071; PubMed-10804147;							
RA	Ando T., Noel J.S., Fankhauser R.L.;							
RT	"Genetic classification of 'Norwalk-like viruses.';							
RL	J. Infect. Dis. 181:S336-S348(2000).							
RN	[3]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=HU/NLV/HONOLULU/219/1992/US;							
RA	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;							
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.							
DR	EMBL; AF414403; AAL12962.1; -							
DR	InterPro; IPR004005; Calici_coat.							
DR	Pfam; PF00915; Calici_coat; 1.							
SQ	SEQUENCE 545 AA; 58631 MW; 89D035D4FEB10BEF CRC64;							
Query Match 98.3%; Score 2846; DB 12; Length 545;								
Best Local Similarity 98.3%; Pred. No. 1.6e-220;								
Matches 536; Conservative 4; Mismatches 5; Indels 0; Gaps 0;								
QY	1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGOVNMIDPWIMNNY	60					
Db	1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGOVNMIDPWIMNNY	60					
QY	61	VQAPQGEFTISPNNTPGDILDLQLGPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFTA	120					
Db	61	VQAPQGEFTISPNNTPGDILDLQLGPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFTA	120					

Query Match 98.3%; Score 2846; DB 12; Length 545;  
Best Local Similarity 98.5%; Pred. No. 1.6e-220;

Matches	537:	Conservative	2:	Mismatches	6:	Indels	0:	Gaps	0:
Qy	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPTISMEP	VAGAAATAAATAGOVNMDPWIMNNY	60					
Db	1	MMASKDAPTNMDGTSGAGQLVPEANTAEPTISMEP	VACAAATAAATAGOVNMDPWIMNNY	60					
Qy	61	VOAQGETTISPNTTGGDILFDLQGLPHLNPFLSHLAQYNGWGNKMKVKVLLAGNAFTA	120						
Db	61	VOAQGETTISPNTTGGDILFDLQGLPHLNPFLSHLAQYNGWGNKMKVKVLLAGNAFTA	120						
Qy	121	GKIIISCIIPPCFAAGNTSIAOATWPHFVIADVRLIEVPLEVDVNRVLFHNNNDNAPTMR	180						
Db	121	GKIIISCIIPPCFAAGNTSIAOATWPHFVIADVRLIEVPLEVDVNRVLFHNNNDNAPTMR	180						
Qy	181	LVCMLYTPLRASGSSSGTDPEVIAAGRLVTCPSPDFSFLFVPPNVEQKTKPFSVPNPLPN	240						
Db	181	LVCMLYTPLRASGSSSGTDPEVIAAGRLVTCPSPDFSLLFLVPPNVEQKTKPFSVPNPLPN	240						
Qy	241	TLSNSRVPSLTKSMVSRDHQOMQVQFNGRVTLDGLOGTTPTSASQLCKTRGVSFHANG	300						
Db	241	TLSNSRVPSLTKSMVSRDHQOMQVQFNGRVALDGLOGTTPTSASQLCKTRGVSFHANG	300						
Qy	301	GNNGYNLTDELGSPYHAPESPAPIGFPDLGECDHWEASPTTQFNTGDVVIKQINVKQESAF	360						
Db	301	GNNGYNLTDELGSPYHAPESPAPIGFPDLGECDHWEASPTTQFNTGDVVIKQINVKQESAF	360						
Qy	361	APHLGTIQADGLSDVSVNTNMIAGLWSPYSDGHRGVDVDPWIPRYGSTLTTEAAQLAPP	420						
Db	361	APHLGTIQADGLSDVSVNTNMIAGLWSPYSDGHRGVDVDPWIPRYGSTLTTEAAQLAPP	420						
Qy	421	IYPGFGCAIVFFMSDPPPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD	480						
Db	421	IYPGFGCAIVFFMSDPPPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHYLDPD	480						
Qy	481	THNLGFEKLYPECFTWCVPNSSGTPGOTPLPINGVFVFSVMSREFYQLKPVCTAGPACRL	540						
Db	481	THNLGFEKLYPECFTWCVPNSSGTPGOTPLPINGVFVFSVMSREFYQLKPVCTAGPACRL	540						
Qy	541	GIRRS 545							
Db	541	GIRRS 545							
RESULT 4									
Q918C2	ID	Q918C2	PRELIMINARY;	PRT;	545	AA.			
AC	Q918C2:								
DT	01-DEC-2001	(TrEMBLrel. 19, Created)							
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)							
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)							
DE	Capsid protein.								
GN	ORF2.								
OS	Norwalk-like virus NLV/Little Rock/316/1994/US.								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;								
OC	Norwalk-like viruses.								
OX	NCBI_TaxID=471837;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;								
RX	MEDLINE=97193806; PubMed=9041391;								
RA	Ando T., Monroe S., Noel J.S., Glass R.I.;								
RT	"A one-tube method of reverse transcription-PCR to efficiently amplify								
RT	a 3-kilobase region from the RNA polymerase gene to the poly(A) tail								
RT	of small round-structured viruses (Norwalk-like viruses).";								
RL	J. Clin. Microbiol. 35:570-577(1997).								
RN	[2]								
RP	SEQUENCE OF 100-192 FROM N.A.								
RC	STRAIN-HU/NLV/LITTLE ROCK/316/1994/US;								

RT	nonbacterial acute gastroenteritis in the United States, 1990 to				
RL	1995.";				
RL	J. Med. Virol. 53:372-383(1997).				
RN	{3}				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;				
RX	MEDLINE=20266071; PubMed=10804147;				
RA	Ando T., Noel J.S., Fankhauser R.L.;				
RA	EMBL: AF414405; AAL12968.1; "				
RL	"Genetic classification of 'Norwalk'-like viruses.";				
RT	J. Infect. Dis. 181:S336-S348(2000).				
RN	{4}				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;				
RA	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;				
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF414405; AAL12968.1; "				
DR	InterPro: IPR004005; Calici_coat.				
DR	Pfam: PF00915; Calici_coat; 1.				
SO	SEQUENCE 545 AA; 58666 MW; F55076D61D612B82 CRC64;				
	Query Match	98.1%;	Score 2840;	DB 12;	Length 545;
	Best Local Similarity	98.5%;	Pred. No. 4.7e-220;		
	Matches 537;	Conservative 3;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1	MMASKADPTNMDCGTSAGQLVPEANTAEPISMEPISBPVAGAATAAATAGOVNMIDPWIMNNY	60		
Db	1	MMASKADAPPNMDGTSAGQLVPEANTAEPISMEPISBPVAGAATAAATAGOVNMIDPWIMNNY	60		
Qy	61	VOAQGETTISPNTTGGDILFDLQGLPHLNPFLSHLAQMYNGWGNKKVYLLAGNAFTA	120		
Db	61	VOAQGETTISPNTTGGDILFDLQGLPHLNPFLSHLAQMYNGWGNKKVYLLAGNAFTA	120		
Qy	121	GKIIISCIPPGFAAQTISIAQATWPHVADRVLEPIEVPLEDVRNVLPHNNNDNPTMR	180		
Db	121	GKIIISCIPPGFAAQTISIAQATWPHVADRVLEPIEVPLEDVRNVLPHNNNDNPTMR	180		
Qy	181	LVCMLYTPLRASGSSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNPLN	240		
Db	181	LVCMLYTPLRASGSSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNPLN	240		
Qy	241	TLNSRVPSLIKSMVSRDHQWQFQNGRYTLDGLOGTTPTSASOLCKIRGSVFIANG	300		
Db	241	TLNSRVPSLIKSMVSRDHQWQFQNGRYTLDGLOGTTPTSASOLCKIRGSVFIANG	300		
Qy	301	NGYNLTDELDCSPYHATESPAPIGFPDLGECDHWEASPTTQFNTGDVVIKQINVKQESAF	360		
Db	301	NGYNLTDELDCSPYHATESPAPIGFPDLGECDHWEASPTTQFNTGDVVIKQINVKQESAF	360		
Qy	361	APHLGTIQAGLSDVSVNTNNIAKLGWVSPVSDGHRGVDVDPWIPRYGSLTEAAQLAPP	420		
Db	361	APHLGTIQAGLSDVSVNTNNIAKLGWVSPVSDGHRGVDVDPWIPRYGSLTEAAQLAPP	420		
Qy	421	IYPGFGGAIYVFMSDPPPIAHGTNGLSVPTCIPQEFVTHFVNEQAPRGEAALHLYLDPD	480		
Db	421	IYPGFGGAIYVFMSDPPPIAHGTNGLSVPTCIPQEFVTHFVNEQAPRGEAALHLYLDPD	480		
Qy	481	THRNLFGEKLYPEGFMTCVPNSSCTGPQTLPIINGVFVSVWSRYQLKPVGTAGPACRL	540		
Db	481	THRNLFGEKLYPEGFMTCVPNSSCTGPQTLPIINGVFVSVWSRYQLKPVGTAGPACRL	540		
Qy	541	GIRRS 545			
Db	541	GIRRS 545			
RESULT 5					
Q91H12	PRELIMINARY; PRT; 545 AA.				
ID	Q91H12				AC
AC	Q91H12;				
DT	01-DEC-2001	(TrEMBLrel. 19, Created)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)			
DE	Capsid.				

OS Human calicivirus NLV/VA98115/1998.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.  
OX NCBI\_TaxID=165509;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/VA98115/1998;  
RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,  
RA Fulton D., Morrow R., Watson D.O.;  
RT "Baculovirus expression and antigenic characterization of the capsid  
RT proteins of three Norwalk-like viruses.";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY038598; AAK84673.1; -;  
DR InterPro: IPR004005; Calici.coat.  
DR Pfam: PF00915; Calici.coat.1;  
SQ SEQUENCE 545 AA; 58295 MW; 259B0C7DEDB34DAC CRC64;

Query Match 96.6%; Score 2798; DB 12; Length 545;  
Best Local Similarity 96.7%; Pred. NO. 1.le-216;  
Matches 527; Conservative 7; Mismatches 11; Indels 0; Gaps 0;  
QY 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60  
Db 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60  
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120  
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120  
QY 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVRNLFHNDNAPTMR 180  
Db 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVRNLFHNDNAPTMR 180  
QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
QY 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300  
Db 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300  
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDSHMEASPTTQFNTGDIKQINVKQESAF 360  
Db 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDSHMEASPTTQFNTGDIKQINVKQESAF 360  
QY 361 APHLGTTQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVPWIPRYGSTLTEAAQLAPP 420  
Db 361 APHLGTTQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVPWIPRYGSTLTEAAQLAPP 420  
QY 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCPTIPEFVTHFVNEQAPTRGEAALLHYLDPD 480  
Db 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCPTIPEFVTHFVNEQAPTRGEAALLHYLDPD 480  
QY 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540  
Db 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540  
QY 541 GIRRS 545  
Db 541 GIRRS 545

RESULT 6  
Q9IV45 PRELIMINARY; PRT; 545 AA.  
ID Q9IV45  
AC Q9IV45;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/VA98115/1998.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.

OX NCBI\_TaxID=122916;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/BIRMINGHAM/93/UK;  
RX MEDLINE=20404883; PubMed=10949950;  
RA Green J., Vinje J., Callimore C.I., Koopmans M., Hale A.D.,  
RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
RT "Capsid protein diversity among 'Norwalk-like' viruses.";  
RL Virus Genes 20:227-236(2000).  
DR EMBL: AJ277612; CAB89093.1; -;  
DR InterPro: IPR004005; Calici.coat.  
DR Pfam: PF00915; Calici.coat.1;  
SQ SEQUENCE 545 AA; 58689 MW; DC6EB06EBECDF7C7 CRC64;

Query Match 94.8%; Score 2744; DB 12; Length 545;  
Best Local Similarity 95.8%; Pred. NO. 2.5e-212;  
Matches 522; Conservative 7; Mismatches 16; Indels 0; Gaps 0;  
QY 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60  
Db 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60  
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120  
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120  
QY 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVRNLFHNDNAPTMR 180  
Db 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVRNLFHNDNAPTMR 180  
QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240  
QY 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300  
Db 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300  
QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDSHMEASPTTQFNTGDIKQINVKQESAF 360  
Db 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDSHMEASPTTQFNTGDIKQINVKQESAF 360  
QY 361 APHLGTTQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVPWIPRYGSTLTEAAQLAPP 420  
Db 361 APHLGTTQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVPWIPRYGSTLTEAAQLAPP 420  
QY 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCPTIPEFVTHFVNEQAPTRGEAALLHYLDPD 480  
Db 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCPTIPEFVTHFVNEQAPTRGEAALLHYLDPD 480  
QY 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540  
Db 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540  
QY 541 GIRRS 545  
Db 541 GIRRS 545

RESULT 7  
Q66418 PRELIMINARY; PRT; 544 AA.  
ID Q66418  
AC Q66418;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Capsid protein.  
OS Desart Shield virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norwalk-like viruses.  
OX NCBI\_TaxID=33755;  
RN [1]  
RP SEQUENCE FROM N.A.

```
RC STRAIN=DSV2395;
RX MEDLINE=94174735; PubMed=8128635;
RA Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
RT "Molecular characterization and expression of the capsid protein of a
RT Norwalk-like virus recovered from a Desert Shield troop with
RT gastroenteritis.";
RL Virology 200:319-325(1994).
DR EMBL: U04469; AAA16285.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58606 MW; 07A4AE7E4BCE81E8 CRC64;

Query Match 90.0%; Score 2605.5; DB 12; Length 544;
Best Local Similarity 89.2%; Pred. No. 3.5e-201;
Matches 486; Conservative 24; Mismatches 30; Indels 1; Gaps 1;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMEPHVIAADVRLVLEPIEVLVDVNRVLFHNNDNAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMEPHVIAADVRLVLEPIEVLVDVNRVLFHNNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLDGLOGQTTPTSASQLCKIRGSVFHANG 300
DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLDGLOGQTTPTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWHM--EASPTTQFNTGDVTKQINVKQESAF 360
DB 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWHM--EASPTTQFNTGDVTKQINVKQESAF 360
QY 361 APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLAPP 420
DB 361 APHLGHVQADNLIS-AGANTDLIVLSWISVSDQHRHDVDPWVPIRYGSSLTEAAQLAPP 419
QY 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPINGVVFVSWVSRYQLKPVGTAGPACRL 540
DB 481 THRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPINGVVFVSWVSRYQLKPVGTAGPACRL 540
QY 541 GIRR 545
DB 540 GIRR 544

RESULT 8
Q9IV48 PRELIMINARY; PRT: 539 AA.
AC Q9IV48:
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Winchester/94/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/WINCHESTER/94/UK;
RX MEDLINE=20404883; PubMed=10949950;
```

```
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RT Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277609; CAB89090.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 1.9e-162;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMEPHVIAADVRLVLEPIEVLVDVNRVLFHNNDNAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMEPHVIAADVRLVLEPIEVLVDVNRVLFHNNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLDGLOGQTTPTSASQLCKIRGSVFHANG 300
DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFNGRVTLDGLOGQTTPTSASQLCKIRGSVFHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWHM--EASPTTQFNTGDVTKQINVKQES 358
DB 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWHM--EASPTTQFNTGDVTKQINVKQES 358
QY 359 AFAPHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLA 418
DB 359 PICTXLGSIET--TSQDQDPAQDGLGTLAWVSPSTSCAR--VDPWKIPSYGSTVTESTILA 414
QY 419 PIYPPGGEAIVFFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLD 478
DB 419 PIYPPGGEAIVFFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLD 478
QY 479 PDTHNLGEFKLYPEGFMTCVPNSGSGTGPQTLPINGVVFVSWVSRYQLKPVGTAGPAC 538
DB 479 PDTHNLGEFKLYPEGFMTCVPNSGSGTGPQTLPINGVVFVSWVSRYQLKPVGTAGPAC 538
QY 539 GLIRR 544
DB 533 GLGVR 538

RESULT 9
Q9DU46 PRELIMINARY; PRT: 544 AA.
AC Q9DU46:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RX Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

Thu Jan 16 15:59:23 2003

```

RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RL expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL: AB042808; BAB18267.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 68.3%; Score 1979; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 9e-151;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPNTMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPNTMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
QY 121 GKVIICVPPGFGQSRITLSIAQAATLFPHVIAADVRLDPEVPLEDVRLVFNHNDTQPTMR 180
DB 121 GKVIICVPPGFGQSRITLSIAQAATLFPHVIAADVRLDPEVPLEDVRLVFNHNDTQPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
QY 241 TLSNSRVPSTLTKSMVSRDHGMQVQNGRVTLTGQLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNSRVPSTLTKSMVSRDHGMQVQNGRVTLTGQLOGTTPTSASOLCKIRGSVFHANG 300
QY 241 YLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
DB 241 YLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
QY 301 GNGYNLTLDGSPHAFESPAPIGFDPDLGECDDHMEAS--PTTQFNTGDVQIKQINVKQES 358
DB 301 GNGYNLTLDGSPHAFESPAPIGFDPDLGECDDHMEAS--PTTQFNTGDVQIKQINVKQES 358
QY 359 AFAPHLGTQADGLSDVSVNTMIATKLGWVSPVSDGHRGVDVWPVPRYGSTLTAQAOLA 418
DB 359 AFAPHLGTQADGLSDVSVNTMIATKLGWVSPVSDGHRGVDVWPVPRYGSTLTAQAOLA 418
QY 359 QFVPHLSITLID--ENVSSGGDYIGTQWTSPPSDSGGANTFWKIPDYGSSLAESAOLA 416
DB 359 QFVPHLSITLID--ENVSSGGDYIGTQWTSPPSDSGGANTFWKIPDYGSSLAESAOLA 416
QY 419 PPIYPGGEAIVFMSDFFPIAHGTNGLS----VPTIPIQEFVTHFVNEQAPTRGEAALL 474
DB 419 PPIYPGGEAIVFMSDFFPIAHGTNGLS----VPTIPIQEFVTHFVNEQAPTRGEAALL 474
QY 417 PAVYPPGNEVIVFMASIP---GPNQSGSNLPVCLLPQEIYTHFISEQAPIQGEAALL 473
DB 417 PAVYPPGNEVIVFMASIP---GPNQSGSNLPVCLLPQEIYTHFISEQAPIQGEAALL 473
QY 475 HYLDPDTHRNLFGEKLYPEGMTCVPNSSGTGPOTLPINGVVFVSWVSRFYQLKPVGTA 534
DB 475 HYLDPDTHRNLFGEKLYPEGMTCVPNSSGTGPOTLPINGVVFVSWVSRFYQLKPVGTA 534
QY 535 GPA-CRLGIRR 544
DB 535 GPA-CRLGIRR 544
QY 534 GPARGRLGVR 544
DB 534 GPARGRLGVR 544

RESULT 10
Q91185 PRELIMINARY; PRT; 544 AA.
AC Q91185;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/KOBLENZ 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated

```

```

RT with Norwalk-like viruses in Germany.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF394960; AAK72048.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A68 CRC64;

Query Match 68.2%; Score 1975; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 1.9e-150;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPNTMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPNTMDGTSGAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
QY 121 GKVIICVPPGFGQSRITLSIAQAATLFPHVIAADVRLDPEVPLEDVRLVFNHNDTQPTMR 180
DB 121 GKVIICVPPGFGQSRITLSIAQAATLFPHVIAADVRLDPEVPLEDVRLVFNHNDTQPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNPLN 240
QY 241 TLSNSRVPSTLTKSMVSRDHGMQVQNGRVTLTGQLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNSRVPSTLTKSMVSRDHGMQVQNGRVTLTGQLOGTTPTSASOLCKIRGSVFHANG 300
QY 241 YLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
DB 241 YLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
QY 301 GNGYNLTLDGSPHAFESPAPIGFDPDLGECDDHMEAS--PTTQFNTGDVQIKQINVKQES 358
DB 301 GNGYNLTLDGSPHAFESPAPIGFDPDLGECDDHMEAS--PTTQFNTGDVQIKQINVKQES 358
QY 359 AFAPHLGTQADGLSDVSVNTMIATKLGWVSPVSDGHRGVDVWPVPRYGSTLTAQAOLA 418
DB 359 AFAPHLGTQADGLSDVSVNTMIATKLGWVSPVSDGHRGVDVWPVPRYGSTLTAQAOLA 418
QY 359 QFVPHLSITLID--DNVSSGGDYIGTQWTSPPSDSGGANTFWKIPDYGSSLAESAOLA 416
DB 359 QFVPHLSITLID--DNVSSGGDYIGTQWTSPPSDSGGANTFWKIPDYGSSLAESAOLA 416
QY 419 PPIYPGGEAIVFMSDFFPIAHGTNGLS----VPTIPIQEFVTHFVNEQAPTRGEAALL 474
DB 419 PPIYPGGEAIVFMSDFFPIAHGTNGLS----VPTIPIQEFVTHFVNEQAPTRGEAALL 474
QY 417 PAVYPPGNEVIVFMASIP---GPNQSGSNLPVCLLPQEIYTHFISEQAPIQGEAALL 473
DB 417 PAVYPPGNEVIVFMASIP---GPNQSGSNLPVCLLPQEIYTHFISEQAPIQGEAALL 473
QY 475 HYLDPDTHRNLFGEKLYPEGMTCVPNSSGTGPOTLPINGVVFVSWVSRFYQLKPVGTA 534
DB 475 HYLDPDTHRNLFGEKLYPEGMTCVPNSSGTGPOTLPINGVVFVSWVSRFYQLKPVGTA 534
QY 535 GPA-CRLGIRR 544
DB 535 GPA-CRLGIRR 544
QY 534 GPARGRLGVR 544
DB 534 GPARGRLGVR 544

RESULT 11
Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHIBA 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a
RT genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB022679; BAA82106.1; -.

```

```
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1;
SQ SEQUENCE 544 AA: 58323 MW: ABD1CLCF4F93D872 CRC64;

Query Match 68.2%; Score 1974; DB 12; Length 544;
Best Local Similarity 66.6%; Pred. No. 2.3e-150;
Matches 367; Conservative 68; Mismatches 102; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MMASKDAPTSADGATGAGQLVPEVNTADPIIDPVAGSALATAGVNLDPWIINF 60
Qy 61 VOAPOGETISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMKVKVLLAGNAFTA 120
Db 61 VOAPOGETISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMKVVRVYLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPTEVPLEDVRNVLFNHNDNAPTMR 180
Db 121 GKVIICVPPGFGOSRTLSIAQATLPHVIADRVTLDPVEVPLEDVRNVLHNNDTQPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPLN 240
Db 181 LLCMLYTPLRGASGSGTDSFVAVAGRVLTCPGPDFLFLVPPTVEQKTRPTVDNIPLK 240
Qy 241 TILNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 YLSNRIENPNEGMSLSDPQTONVQFQNGRGTIDGQLGTTVPVSYLSOLCKERGRI--TSG 298
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPDLGECGDWHMEAS--PTTQENTGDVTKQINVKQES 358
Db 299 QRVNLNLTLDGSPFMGFCAPAPAGPDLGSCDWHIEMSKIPNSSTQNNPIVNSVKPNSQ 358
Qy 359 AFAPHLGTQADGLSDVSVNTNMIAKLQWSPVSDGHRGVDPPWIPRYGSLTLEAAQLA 418
Db 359 QFVPHLSSTILD--ENVSSGGYIGTIQWTSPDSDGGANTNFKWIPDYGGSLAEASQLA 416
Qy 419 PPIYPGGEAIVFMSDFPTAHGTNGLS---VPTTPOEFVTHFVNEQAPTRGEAALL 474
Db 417 FAVYPPGNEVIVFMAIP---GNQSGSNLVPCLLPQEIYTHIFSEQAPIQGEAALL 473
Qy 475 HYLDPDTHNLGEFLKYPGEFMTCPVNSSGTGPTPLPINGVFVFSWVSRYQLKPVGTA 534
Db 474 HYVDPDTHNLGEFLKYPGGLYTCVPSNSSGTGPQPLDGVFVFAVSRYQLKPVGTA 533
Qy 535 GPA-CRLGIRR 544
Db 534 GPARGRLGVR 544

RESULT 12
Q9YS14 PRELIMINARY; PRT; 540 AA.
AC Q9YS14;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWALK-LIKE;
RX MEDLINE=20254531; PubMed=10795514;
RA Schreier E., Doering F., Kuenkel U.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with small round structured viruses in Germany in 1997/98.";
RL Arch. Virol. 145:443-453(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWALK-LIKE;
RA Schreier E., Doering F., Kuenkel U.;
```

```
RL Submitted (SEP-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL: AF093797; AAC64603.1;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1;
SQ SEQUENCE 540 AA: 58018 MW: 647843487A654CBE CRC64;

Query Match 68.1%; Score 1972.5; DB 12; Length 540;
Best Local Similarity 67.7%; Pred. No. 3e-150;
Matches 377; Conservative 60; Mismatches 89; Indels 31; Gaps 7;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MMASKDAPPSDAGSAGQLVPEVNTADQISMVPVAGATAVATAGVNMIDPWIFNMF 60
Qy 61 VOAPOGETISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMKVKVLLAGNAFTA 120
Db 61 VOAPOGETISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMKVVRVILLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPTEVPLEDVRNVLFNHNDNAPTMR 179
Db 121 GKVIICVPPGFGDARILTAQATLPHVIADRVTLDPVEVPLEDVRNVLHNSSQPQPTM 180
Qy 180 RLVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFSVPLNPL 239
Db 181 RUVAMLYTPLRGSGSGTDAFVAVAGRVLTCPAPDFSLFLVPPSVQEKTRVFSVPNIPL 240
Qy 240 NTLNSRVPSLIKSMVSRDHQGMVQFQNGRVTLDCQLQGTTPTSASQLCKIRGSVFHAN 299
Db 241 KDLNSRVPTLQGMFVSPDVNQSVQFQNGRCQIDQLGTTVPVLSQLCKIRGRT--SS 298
Qy 300 GNGYNLTLDGSPYHAFESPAPIGPDLGECGDWHM-----EASPTTQNTGVDVTK 350
Db 299 NTRVLNLTSEVDGTPVPLESPAPVPGFPGDGGDWHVGETFEARDQDPQSONVTEATND--- 355
Qy 351 QINVQESAFAPHLGTI--QADGLSDVSVNTNMIAKLQWSPVSDGHRGVDPPWIPRY 407
Db 356 -----SSEVYLGSI SPHNGDGFH---SCDIIGSLDMSAPSDGSALDV--WSIPKY 402
Qy 408 GSTLEAQAQAPPIYPGGEAIVFMSDFPTAHGTNGLSVPCCTIPQEFVTHFVNEQAPT 467
Db 403 GSSLPDVTHLAPAVPPGFGVEVILYFHSKFGSGPTDKLRVPCLLIPQEIYTHIFCNEQAPI 462
Qy 468 RGEAALLHYLDPDTHNLGEFLKYPGEFMTCPVNSSGTGPTPLPINGVFVFSWVSRYQLK 527
Db 463 AGEAALLHYVDPDTHNLGEFLKYPGEFMTCPVNSISSGPTPLPINGVFVFSWVSRYQLK 522
Qy 528 LKPVGTAGPACRLGIRR 544
Db 523 LKPVGTASAARLGIUR 539

RESULT 13
Q918C5 PRELIMINARY; PRT; 544 AA.
AC Q918C5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus NLV/Baltimore/277/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=171836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
```



RESULT 14	
Q9IV47	PRELIMINARY: PRT: 546 AA.
AC Q9IV47	
DT 01-OCT-2000 (TrEMBLrel. 15, Created)	
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE Capsid protein.	
OS Human calicivirus HU/NLV/Whiterose/96/UK.	
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	
OC Norwalk-like viruses.	
OX NCBI_TaxID=122914;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=HU/NLV/WHITEROSE/96/UK.	
RX MEDLINE=20404883; PubMed=1094950;	
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,	
RA Clegg J.C., Chamberlain J., Brown D.W.G.;	
RT "Capsid protein diversity among 'Norwalk-like' viruses.";	
RL Virus Genes 20:227-236(2000).	
DR EMBL; AJ277610; CAB85091.1; .	
DR InterPro; IPR004005; Calici_coat.	
DR Pfam; PF00915; Calici_coat; 1.	
SQ SEQUENCE 546 AA; 58737 MW; 501F9D2BE2A77B51 CRC64;	
Query Match 67.8%; Score 1964.5; DB 12; Length 546;	
Best Local Similarity 66.2%; Pred. No. 1.3e-149;	
Matches 365; Conservative 78; Mismatches 95; Indels 13; Gaps 7;	
Qy 1 MMASKDAPNTMDGTSCAGQLVPEANTAEPISEPVAGATAAATAGOVNMIDPWIMNY 60	
Db 1 MMASKGAPQSGAGSGAGQLVPEANTADPLMEPVAGPTTAVATAGOVNMIDPWIMNF 60	
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120	
Db 61 VQSPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNRVRIILLAGNAFSA 120	
Qy 121 GKTIISCIIPGFAAONISIAQATMFPHVIADRVLPVLEPIEVPLEDVRNVLFHNNDAPTMR 180	
Db 121 GKIIICVCPGFTSSSLTIAQATLFPHVIADRVLTLEPIEMPLEDVRNVLHYTNDSTQPTMR 180	
Qy 181 LVCMLYPLRASGSSSGTDFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSPVNLPLN 240	
Db 181 LVCMLYPLRTGGSGNSDSFVAGRVLTAPSDSFLFLVPTIEQKTRAEVTPNIPLQ 240	
Qy 241 TILNSRVPSLTKSMVSRDHQMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVFHANG 300	
Db 241 TILNSRFPFLQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVGRKI--NQG 298	
Qy 301 GNGYNLTLDGSPHAFESPAPIGPDGLGCDWHMEASPT--TOFNTGDVVKQINVKQES 358	
Db 299 ARTLNLTEVDGKPFMAFDSPAPVGFDPFGKCDWHIRVSKTPNNTSSGDPMSRSVSVQINVO 358	
Qy 359 AFAPHLGTIQADGLSDSVNTNMTAKLQWSPVSDGHRGVDVDPWVIPRYGSTLTEAAQLA 418	
Db 359 GVPVHLGSLQFDEVENHPGTG-DYICTIEMISQSPSTPLCTDNLWEIPDYGSSLSQAANLA 417	
Qy 419 PPIYPGGEAIVFMSDFPIAHGTNGLS----VPCITPQEFVTHFVNEQAPTGEAALL 474	
Db 418 PVVFPFGGEALYFVSFAFP--GPNRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALL 474	
Qy 475 HYLDPDTHNLGFEKLYPEGEFTCPVNSSGTCGPOTPLINGVFVSVWSRYQLKPVGTA 534	
Db 475 HVYDDPTNRLNGFEKLYPGGYLTCVPGNGVAGAPQOLPLNGVFLFVSVWSRYQLKPVGTA 534	
Qy 535 GPA-CRLGIRR 544	
Db 535 STARGRLGVR 545	
RESULT 15	
Q9IV43	PRELIMINARY: PRT: 543 AA.
ID Q9IV43	



```
AC O9IV43:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Musgrove/89/UK.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122918;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/MUSGROVE/89/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277614; CAB89095.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58216 MW; 11F97A3989CC540 CRC64;

Query Match 67.6%; Score 1956.5; DB 12; Length 543;
Best Local Similarity 67.8%; Pred. No. 5.8e-149;
Matches 377; Conservative 65; Mismatches 89; Indels 25; Gaps 10;

QY 1 MMASKDAPTINMDGTSGAGOLVPEANTAEPISMPEVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDATPSADGAGOLVPEVNNAXPLPLDPVAGASTALATAGQVNMIDPWIFNNF 60
QY 61 VQAPQGEFTISPNTGDIILFDLQPLHNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTGDIILFDLQPLHNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GTIISCIPPGFAAQNISAQATMPEHVIADRVLEPIEVPLEDVRLVLFHNNNDAPTMR 180
DB 121 GKVIICVPPGFSRLSTAQATLFFHIADVRTLEPIELPVEDVRLTYHTNDNQPTMR 180
QY 181 LVCMLYTLPRAGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBQKTKPFSVPLNPLN 240
DB 181 LLCLMLYTLRTGGGGTDAFVAVAGRVLTCPSSDFNLFLLVPTVEQKTRPFSVPIPLQ 240
QY *241 TLSNRVPSLIKSMVSRDHGQWVQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 LLSNRVPLNIQSMVLSPOAQWVQNGRCTTDGOLLGTPVSVSQILKFRGKV--SAG 298
QY 301 GNGYNLTLDGSPYHAFESPAPITGFPLGECDDHMEAS-PTTQFNTGDVIKQINVKQESA 359
DB 299 SKVINLTLDGSPFLAFEPAPPTGFPDLGTSWDHIEMSLNSNSQSSGNPILLRDIQPNSS 358
QY 360 -FAPHLGTTQADGLSDVSVNTNNIAKLGW-----VSPVSDGHRGDVDPWVPIRYGSTLT 412
DB 359 DFVPHLGVSVTTAIDTAGDT--LGTIQWTSQPSNVTVP-----DYNFTIPIQYSSLA 411
QY 413 EAAQLAPPYPPGFGGAIVFFMSDFP---IAHGTNGLSVPTCIPOEFVTHFVNEQAPTRG 469
DB 412 EASQLAPVVPYPPGFGGAIVFFMSPIGPNHTAKPN--LVPCLLPQEFVTHFVSEQAPSMG 469
QY 470 EALLHYLDPDTHRNIGERKLYPEGFMTCVPNSSGTGPGQTLPIGVGVFVSWVSREYQLK 529
DB 470 EAAHYVDPDTHRNIGERKLYPEGFITCPVN--GTGPOOLPLNGVVFVSWVSREYQLK 527
QY 530 PVGTAGPA-CRLGIRR 544
DB 528 PVGTASSARGRLGVR 543
```

Search completed: January 16, 2003, 09:55:56  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**